

531 Rec'd PC. 13 SEP 2001

FORM PTO-1390 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER 4239-60771
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. § 371		
INTERNATIONAL APPLICATION NO. PCT/US00/06946	INTERNATIONAL FILING DATE 16 March 2000	US APPLICATION NO. (if known) <small>37 C.F.R. § 1.15</small> 09/936702
TITLE OF INVENTION A NOVEL CHIMERIC PROTEIN FOR PREVENTION AND TREATMENT OF HIV INFECTION		
APPLICANT(S) FOR DO/EO/US Edward A. Berger, Christie M. Del Castillo		
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:		
<ol style="list-style-type: none"> <li><input checked="" type="checkbox"/> This is a <b>FIRST</b> submission of items concerning a filing under 35 U.S.C. § 371.</li> <li><input type="checkbox"/> This is a <b>SECOND</b> or <b>SUBSEQUENT</b> submission of items concerning a filing under 35 U.S.C. § 371.</li> <li><input type="checkbox"/> This is an express request to begin national examination procedures (35 U.S.C. § 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. § 371(b) and PCT Articles 22 and 39(1).</li> <li><input checked="" type="checkbox"/> A proper Demand for International Preliminary Examination was made by the 19<sup>th</sup> month from the earliest claimed priority date.</li> <li><input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. § 371(c)(2))           <ol style="list-style-type: none"> <li><input type="checkbox"/> is transmitted herewith (required only if not transmitted by the International Bureau).</li> <li><input type="checkbox"/> has been transmitted by the International Bureau.</li> <li><input checked="" type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US).</li> </ol> </li> <li><input type="checkbox"/> A translation of the International Application into English (35 U.S.C. § 371(c)(2)).</li> <li><input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. § 371(c)(3))           <ol style="list-style-type: none"> <li><input type="checkbox"/> are transmitted herewith (required only if not transmitted by the International Bureau)</li> <li><input type="checkbox"/> have been transmitted by the International Bureau.</li> <li><input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired.</li> <li><input checked="" type="checkbox"/> have not been made and will not be made.</li> </ol> </li> <li><input type="checkbox"/> A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. § 371(c)(3)).</li> <li><input checked="" type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. § 371(c)(4)).</li> <li><input type="checkbox"/> A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. § 371(c)(5))</li> </ol>		
Items 11. to 16. below concern document(s) or information included:		
<ol style="list-style-type: none"> <li><input type="checkbox"/> An Information Disclosure Statement under 37 C.F.R. §§ 1.97 and 1.98, PTO1449, copies of references cited (56).</li> <li><input checked="" type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 C.F.R. §§ 3.28 and 3.31 and the Recordal fee of \$40.00 is included.</li> <li><input checked="" type="checkbox"/> A <b>FIRST</b> preliminary amendment.           <ol style="list-style-type: none"> <li><input type="checkbox"/> A <b>SECOND</b> or <b>SUBSEQUENT</b> preliminary amendment.</li> </ol> </li> <li><input type="checkbox"/> A substitute specification.</li> <li><input type="checkbox"/> A change of power of attorney and/or address letter.</li> <li><input checked="" type="checkbox"/> Sequence Listing:           <ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> Paper copy (6 pages)</li> <li><input checked="" type="checkbox"/> Computer-readable form.</li> <li><input checked="" type="checkbox"/> Statement in Compliance</li> </ul> </li> </ol>		
		 <b>24197</b>

531 Rec'd PC

13 SEP 2001

U.S. APPLICATION NO. (If known) (37 C.F.R. § 1.18)	INTERNATIONAL APPLICATION NO	ATTORNEY'S DOCKET NUMBER
097936702	PCT/US00/06946	4239-60771
17. <input checked="" type="checkbox"/> The following fees are submitted:		CALCULATIONS (PTO USE ONLY)
<b>BASIC NATIONAL FEE (37 C.F.R. §§ 1.492(a)(1)-(5):</b>		
Neither International Preliminary Examination fee (37 C.F.R. § 1.482) nor International Search fee (37 C.F.R. § 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO. \$1,000.00		
International Preliminary Examination fee (37 C.F.R. § 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO. \$860.00		
International Preliminary Examination fee (37 C.F.R. § 1.482) not paid to USPTO but International Search fee (37 C.F.R. § 1.445(a)(2)) paid to USPTO as an International Searching Authority. \$710.00		
International Preliminary Examination fee paid to USPTO (37 C.F.R. § 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4). \$690.00		
International Preliminary Examination fee paid to USPTO (37 C.F.R. § 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4). \$100.00		
<b>ENTER APPROPRIATE BASIC FEE AMOUNT =</b> <b>\$ 860.00</b>		
Surcharge of \$130.00 for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 C.F.R. § 1.492(e)).		
CLAIMS	NUMBER FILED	NUMBER EXTRA
Total claims	55 - 20 =	35
Independent Claims	3 - 3 =	0
<b>MULTIPLE DEPENDENT CLAIM(S) (if applicable)</b>		+ \$270.00
<b>TOTAL OF ABOVE CALCULATIONS =</b> <b>\$ 1,490.00</b>		
<input type="checkbox"/> Reduction of 1/2 for filing by small entity. Small entity status is claimed for this application.		
<b>SUBTOTAL =</b> <b>\$ 1,490.00</b>		
Processing fee of \$130.00 for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 C.F.R. § 1.492(f)).		
<b>TOTAL NATIONAL FEE =</b> <b>\$ 1,490.00</b>		
Fee for recording the enclosed assignment (37 C.F.R. § 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 C.F.R. §§ 3.28, 3.31). \$40.00 per property.		
<b>TOTAL FEES ENCLOSED =</b> <b>\$ 1,530.00</b>		
		REFUND → \$
		CHARGE → \$

a.  A check in the amount of \$ 1,530.00 to cover the above fees is enclosed.  
 b.  Please charge my Deposit Account No. \_\_\_\_\_ in the amount of \$ \_\_\_\_\_ to cover the above fees. A duplicate copy of this sheet is enclosed.  
 c.  The Director is hereby authorized to charge any additional fees that may be required, or credit any overpayment, to Deposit Account No. 02-4550. A duplicate copy of this sheet is enclosed.  
 d.  Please return the enclosed postcard to confirm that the items listed above have been received.

**NOTE:** Where an appropriate time limit under 37 C.F.R. § 1.494 or § 1.495 has not been met, a petition to revive (37 C.F.R. § 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

KLARQUIST SPARKMAN, LLP  
One World Trade Center, Suite 1600  
121 S.W. Salmon Street  
Portland, OR 97204-2988

SIGNATURE  
 Tanya M. Harding, Ph.D.  
 NAME  
 42,630  
 REGISTRATION NUMBER

cc: Docketing

531 Rec'd PCT 13 SEP 2001  
IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Berger and Del Castillo Art Unit: Not yet assigned

Application No. Not yet assigned

Filed: Herwith

For: A NOVEL CHIMERIC PROTEIN FOR  
PREVENTION AND TREATMENT OF HIV  
INFECTION

Examiner: Not yet assigned

Date: September 13, 2001

BOX PCT  
COMMISSIONER FOR PATENTS  
WASHINGTON, D.C. 20231

**PRELIMINARY AMENDMENT**

In the Specification, please insert the following header and paragraph on page 1, immediately following the title:

--CROSS REFERENCE TO RELATED APPLICATIONS

This is the National Stage of International Application No. PCT/US00/06946, filed March 16, 2000, and claims the benefit of U.S. Provisional Application No. 60/124,681, filed March 16, 1999. The provisional application is incorporated herein in its entirety.--

By this amendment the specification has been changed to reflect prior related applications. No new matter is added by this amendment.

Respectfully submitted,  
KLARQUIST SPARKMAN, LLP

By

Tanya M. Harding, Ph.D.  
Registration No. 42,630

One World Trade Center, Suite 1600  
121 S.W. Salmon Street  
Portland, Oregon 97204  
Telephone: (503) 226-7391  
Facsimile: (503) 228-9446

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: Berger and Del Castillo

Art Unit: Not yet assigned

Application No. Not yet assigned

Filed: Herewith

For: A NOVEL CHIMERIC PROTEIN FOR  
PREVENTION AND TREATMENT OF  
HIV INFECTION

Examiner: Not yet assigned

Date: September 13, 2001

**STATEMENT IN COMPLIANCE WITH 37 C.F.R. § 1.821(f)**

BOX PCT  
COMMISSIONER FOR PATENTS  
Washington, DC 20231

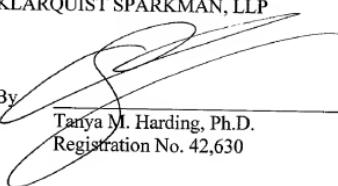
Sir:

In compliance with 37 C.F.R. § 1.821(f), the undersigned declares that the nucleotide and/or amino acid sequences presented in the paper copy of the "Sequence Listing" submitted herewith are the same as the sequences contained in the computer-readable form of the "Sequence Listing."

Respectfully submitted,

KLARQUIST SPARKMAN, LLP

By

  
Tanya M. Harding, Ph.D.  
Registration No. 42,630

One World Trade Center, Suite 1600  
121 S.W. Salmon Street  
Portland, Oregon 97204  
Telephone: (503) 226-7391  
Facsimile: (503) 228-9446

SEQUENCE LISTING

<110> Berger, et al.

<120> NOVEL CHIMERIC PROTEIN FOR PREVENTION AND TREATMENT OF  
HIV INFECTION

<130> 4239-60771

<140>

<141>

<150> PCT/US00/06946

<151> 2000-03-16

<150> 60/124,681

<151> 1999-03-16

<160> 11

<170> PatentIn Ver. 2.1

<210> 1

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: linker peptide

<400> 1'

Gly Gly Gly Gly Ser  
1 5

<210> 2

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: seven repeat  
polypeptide linker

<400> 2

Gly Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly  
1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly  
20 25 30

Gly Gly Ser  
35

<210> 3

<211> 508

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: CD4-scFv(17b)

<400> 3

Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln Leu  
1 5 10 15

Ala Leu Leu Pro Ala Ala Thr Gln Gly Lys Lys Val Val Leu Gly Lys  
20 25 30

Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys Ser  
35 40 45

Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys Ile Leu Gly Asn  
50 55 60

Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys Leu Asn Asp Arg Ala  
65 70 75 80

Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile Ile  
85 90 95

Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr Ile Cys Glu Val Glu  
100 105 110

Asp Gln Lys Glu Glu Val Gln Leu Leu Val Phe Gly Leu Thr Ala Asn  
115 120 125

Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu  
130 135 140

Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly  
145 150 155 160

Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu  
165 170 175

Gln Asp Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys  
180 185 190

Val Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser  
195 200 205

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly  
210 215 220

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly  
225 230 235 240

Gly Gly Ser Gln Val Gln Leu Leu Glu Ser Gly Ala Glu Val Lys Lys  
245 250 255

Pro Gly Ser Ser Val Lys Val Ser Cys Lys Ala Ser Gly Asp Thr Phe  
260 265 270

Ile	Arg	Tyr	Ser	Phe	Thr	Trp	Val	Arg	Gln	Ala	Pro	Gly	Gln	Gly	Leu
275						280						285			
Glw	Trp	Met	Gly	Arg	Ile	Ile	Thr	Ile	Leu	Asp	Val	Ala	His	Tyr	Ala
290					295				300						
Pro	His	Leu	Gln	Gly	Arg	Val	Thr	Ile	Thr	Ala	Asp	Lys	Ser	Thr	Ser
305					310				315			320			
Thr	Val	Tyr	Leu	Glu	Leu	Arg	Asn	Leu	Arg	Ser	Asp	Asp	Thr	Ala	Val
325					330				335						
Tyr	Phe	Cys	Ala	Gly	Val	Tyr	Glu	Gly	Glu	Ala	Asp	Glu	Gly	Glu	Tyr
340					345				350			350			
Asp	Asn	Asn	Gly	Phe	Leu	Lys	His	Trp	Gly	Gln	Gly	Thr	Leu	Val	Thr
355					360				365			365			
Val	Thr	Ser	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly
370					375				380						
Gly	Ser	Glu	Leu	Glu	Leu	Thr	Gln	Ser	Pro	Ala	Thr	Leu	Ser	Val	Ser
385					390				395			400			
Pro	Gly	Glu	Arg	Ala	Thr	Leu	Ser	Cys	Arg	Ala	Ser	Glu	Ser	Val	Ser
	405					410				415					
Ser	Asp	Leu	Ala	Trp	Tyr	Gln	Gln	Lys	Pro	Gly	Gln	Ala	Pro	Arg	Leu
	420					425				430					
Leu	Ile	Tyr	Gly	Ala	Ser	Thr	Arg	Ala	Thr	Gly	Val	Pro	Ala	Arg	Phe
	435					440				445					
Ser	Gly	Ser	Gly	Ala	Glu	Phe	Thr	Leu	Thr	Ile	Ser	Ser	Leu		
	450				455				460						
Gln	Ser	Glu	Asp	Phe	Ala	Val	Tyr	Tyr	Cys	Gln	Gln	Tyr	Asn	Asn	Trp
465					470				475			480			
Pro	Pro	Arg	Tyr	Thr	Phe	Gly	Gln	Gly	Thr	Arg	Leu	Glu	Ile	Lys	Leu
	485					490				495					
Val	Pro	Arg	Gly	Ser	Gly	His	His	His	His	His	His				
	500					505									

<210> 4  
 <211> 1440  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: CD4-scFv(17b)

<400> 4  
 atgaaccggg gagtcccttt taggcacttg ctctgggtc tgcaactggc gtcctccca 60

gcagccactc agggaaagaa agtggtgctg ggcaaaaaag gggatacagt ggaactgacc 120  
tgcacgctt cccagaagaa gaggatcaa ttccacttga aaaactccaa ccaagataaag 180  
attctggaa atcagggtc cttttaact aaagggtccat ccaagctgaa tgcgtcg 240  
gactcaagaa gaagcccttgg gaccaagga aacttcccccc tgcgtccaa gaatcttaag 300  
atagaagact cagatcttgcatctgtac ttggaggaggaa agaaggaggaa ggtgcattg 360  
ctatgttgc tggactgc caactgtac acccacatgc ttccaggggca ggcgtgacc 420  
ctgaccccttgg agagcccccc ttgttagtc ccctcgtgc aatgttaggg tccaaagggtt 480  
aaaaacatc acggggggaa gacccttcc ttggactcgc tggagctcca ggatagtggc 540  
acctggacat gcaactgttgc ccaacccatc aagaagggtt gttccaaat agacatcg 600  
gtgttagtcc tccagaaggc ttccggagggtt ggcgtgttgc tgggggggggg 660  
gtgttagtcc ttggccgggg gtggccgggg ggtgttgc ggggtggggg 720  
ggagggttgc aggttgcgttgc gtcgtatc ggggtgggg tgaagaaggcc tgggttctcg 780  
gtgaagggtt ctgtcaaggc ctctggagac accttcatc gatatagtt tacctgggtt 840  
cgacacggcc ctggacaaaggc ctttggatggg atggaaaggaa tcatcactat ctttgatgtt 900  
gcacactcggccacccatc ccaaggcaga gtcacgatca cccggacaa gtcacccagg 960  
acagtctacc ttggactgtcga gaatctaaatc ttcgtacgatca cccggctata ttctgttgc 1020  
ggagggttgc agggagggc ggacggggggg gaatatgatataatgggtt tctgaacat 1080  
ttggccggggaa gaaccatctgttgc acgggttgc ttcgggttgc ggtgttgc ggggtgggg 1140  
acgggttgc ggggttgc acgggttgc acgggttgc acgggttgc acgggttgc acgggttgc 1200  
ccaggggaa gagccacccatc ctctggaggcc gtcgtgttgc ggtgttagtag cgtacttagcc 1260  
ttgttaccaggc agaaacatccatc ccaagggttgc acgggttgc ttcgggttgc ggtgttgc 1320  
ggccacccatc ttccagggttgc gtcgtgttgc acgggttgc gggcagaattt cacttcacc 1380  
atcggccatc ttcgggttgc acgggttgc acgggttgc acgggttgc acgggttgc 1440

<210> 5  
<211> 127  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: synthetic  
oligonucleotide

<400> 5  
cctccggagg tggcggtagt gggggggggc gttcaggccg aggtggatcc ggaggccgg 60  
ggtcggccgg ggttggaaatc gggggggggc gttctgttgc cggagggttacc actatgttaag 120  
tgtagtag 127

<210> 6  
<211> 39  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: peptide  
encoded by SEQ ID NO: 5

<400> 6  
Ala Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Gly  
1 5 10 15  
Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Gly Ser  
20 25 30  
Gly Gly Gly Gly Thr Thr Ser

<210> 7  
<211> 103  
<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 7

ttttatggat ccgggtggcg agggtcgggc ggggggtggaa ggggggtgg cggctccgga 60  
ggcgaggat cacagggtca gctgtctcgag tctggggctg agg 103

<210> 8

<211> 32

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide  
encoded for by SEQ ID NO: 7

<400> 8

Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Gly  
1 5 10 15

Ser Gly Gly Gly Ser Gln Val Gln Leu Leu Glu Ser Gly Ala Glu  
20 25 30

<210> 9

<211> 65

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 9

taattttatcg atcacgtgac tagtcctagg cccgggtcaa tggatgtatgat gatgtatggcc 60  
gctgc 65

<210> 10

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide  
encoded for by SEQ ID NO: 9

<400> 10  
Ser Gly His His His His His His  
1 5

<210> 11  
<211> 131  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: reverse  
oligonucleotide

<400> 11  
ctagtaactc acttaactag tggtaacctcc gccacctgag ccggccacccc 60  
ccccggccga ccctccggct ccggatccac ctccgcctga accggctccc cactaccggcc 120  
acctccggag g 131

TTGCTTGGG-GGCGGGGGG

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/936,702

DATE: 10/04/2001

TIME: 14:41:19

Input Set : A:\60771 seq.txt

Output Set: N:\CRF3\10042001\I936702.raw

3 <110> APPLICANT: Berger, et al.  
5 <120> TITLE OF INVENTION: NOVEL CHIMERIC PROTEIN FOR PREVENTION AND TREATMENT OF  
6 HIV INFECTION  
8 <130> FILE REFERENCE: 4239-60771  
C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/936,702  
C--> 11 <141> CURRENT FILING DATE: 2001-09-13  
13 <150> PRIOR APPLICATION NUMBER: PCT/US00/06946  
14 <151> PRIOR FILING DATE: 2000-03-16  
16 <150> PRIOR APPLICATION NUMBER: 60/124,681  
17 <151> PRIOR FILING DATE: 1999-03-16  
19 <160> NUMBER OF SEQ ID NOS: 11  
21 <170> SOFTWARE: PatentIn Ver. 2.1  
23 <210> SEQ ID NO: 1  
24 <211> LENGTH: 5  
25 <212> TYPE: PRT  
26 <213> ORGANISM: Artificial Sequence  
28 <220> FEATURE:  
29 <223> OTHER INFORMATION: Description of Artificial Sequence: linker peptide ✓  
31 <400> SEQUENCE: 1  
32 Gly Gly Gly Gly Ser  
33 1 5  
36 <210> SEQ ID NO: 2  
37 <211> LENGTH: 35  
38 <212> TYPE: PRT  
39 <213> ORGANISM: Artificial Sequence  
41 <220> FEATURE:  
42 <223> OTHER INFORMATION: Description of Artificial Sequence: seven repeat ✓  
43 polypeptide linker  
45 <400> SEQUENCE: 2  
46 Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly  
47 1 5 10 15  
49 Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly  
50 20 25 30  
52 Gly Gly Ser  
53 35  
56 <210> SEQ ID NO: 3  
57 <211> LENGTH: 508  
58 <212> TYPE: PRT  
59 <213> ORGANISM: Artificial Sequence  
61 <220> FEATURE:  
62 <223> OTHER INFORMATION: Description of Artificial Sequence: CD4-scFv(17b)✓  
64 <400> SEQUENCE: 3  
65 Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln Leu  
66 1 5 10 15  
68 Ala Leu Leu Pro Ala Ala Thr Gln Gly Lys Lys Val Val Leu Gly Lys  
69 20 25 30  
71 Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys Ser

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/936,702

DATE: 10/04/2001

TIME: 14:41:19

Input Set : A:\60771 seq.txt

Output Set: N:\CRF3\10042001\I936702.raw

72	35	40	45	
74	Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys Ile Leu Gly Asn			
75	50	55	60	
77	Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys Leu Asn Asp Arg Ala			
78	65	70	75	80
80	Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile Ile			
81	85	90	95	
83	Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr Ile Cys Glu Val Glu			
84	100	105	110	
86	Asp Gln Lys Glu Glu Val Gln Leu Leu Val Phe Gly Leu Thr Ala Asn			
87	115	120	125	
89	Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu			
90	130	135	140	
92	Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly			
93	145	150	155	160
95	Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu			
96	165	170	175	
98	Gln Asp Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys			
99	180	185	190	
101	Val Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser			
102	195	200	205	
104	Gly Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly			
105	210	215	220	
107	Gly Gly Gly Ser Gly Gly Ser Gly Gly Gly Ser Gly Gly			
108	225	230	235	240
110	Gly Gly Ser Gln Val Gln Leu Leu Glu Ser Gly Ala Glu Val Lys Lys			
111	245	250	255	
113	Pro Gly Ser Ser Val Lys Val Ser Cys Lys Ala Ser Gly Asp Thr Phe			
114	260	265	270	
116	Ile Arg Tyr Ser Phe Thr Trp Val Arg Gln Ala Pro Gly Gln Gly Leu			
117	275	280	285	
119	Glu Trp Met Gly Arg Ile Ile Thr Ile Leu Asp Val Ala His Tyr Ala			
120	290	295	300	
122	Pro His Leu Gln Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser			
123	305	310	315	320
125	Thr Val Tyr Leu Glu Leu Arg Asn Leu Arg Ser Asp Asp Thr Ala Val			
126	325	330	335	
128	Tyr Phe Cys Ala Gly Val Tyr Glu Gly Glu Ala Asp Glu Gly Glu Tyr			
129	340	345	350	
131	Asp Asn Asn Gly Phe Leu Lys His Trp Gly Gln Gly Thr Leu Val Thr			
132	355	360	365	
134	Val Thr Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly			
135	370	375	380	
137	Gly Ser Glu Leu Glu Leu Thr Gln Ser Pro Ala Thr Leu Ser Val Ser			
138	385	390	395	400
140	Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Glu Ser Val Ser			
141	405	410	415	
143	Ser Asp Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu			
144	420	425	430	



## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/936,702

DATE: 10/04/2001

TIME: 14:41:19

Input Set : A:\60771 seq.txt

Output Set: N:\CRF3\10042001\I936702.raw

207 cctccggagg tggcggtagt gggggaggcg gttcaggcg aggtggatcc ggaggcgag 60  
 208 ggtcggccgg ggttggaaac ggggggtggcg gctctgggg cggaggtacc actatgttaag 120  
 209 tggatag 127

212 <210> SEQ ID NO: 6  
 213 <211> LENGTH: 39  
 214 <212> TYPE: PRT  
 215 <213> ORGANISM: Artificial Sequence  
 217 <220> FEATURE:  
 218 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide ✓  
 219 encoded by SEQ ID NO: 5

221 <400> SEQUENCE: 6

222 Ala Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Gly 15  
 223 1 5 10 15  
 225 Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser 30  
 226 20 25 30  
 228 Gly Gly Gly Gly Thr Thr Ser  
 229 35

232 <210> SEQ ID NO: 7  
 233 <211> LENGTH: 103  
 234 <212> TYPE: DNA  
 235 <213> ORGANISM: Artificial Sequence ✓  
 237 <220> FEATURE:  
 238 <223> OTHER INFORMATION: Description of Artificial Sequence: primer ✓  
 240 <400> SEQUENCE: 7

241 ttttatggat ccgttggcgag ggggtcgccg ggggggtggaa ggggggggtgg cggctccgga 60  
 242 ggccggagggtt cacacgtca gctgctcgag tctggggctg agg 103  
 245 <210> SEQ ID NO: 8  
 246 <211> LENGTH: 32  
 247 <212> TYPE: PRT  
 248 <213> ORGANISM: Artificial Sequence ✓  
 250 <220> FEATURE:  
 251 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide ✓  
 252 encoded for by SEQ ID NO: 7

254 <400> SEQUENCE: 8

255 Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly 15  
 256 1 5 10 15  
 258 Ser Gly Gly Gly Ser Gln Val Gln Leu Leu Glu Ser Gly Ala Glu  
 259 20 25 30

265 <210> SEQ ID NO: 9  
 266 <211> LENGTH: 65  
 267 <212> TYPE: DNA  
 268 <213> ORGANISM: Artificial Sequence ✓  
 270 <220> FEATURE:  
 271 <223> OTHER INFORMATION: Description of Artificial Sequence: primer ✓  
 273 <400> SEQUENCE: 9

274 taattttatcg atcacgtgac tagtcctagg cccgggtcaa tggatgtatgat gatgtatggcc 60  
 275 gctgc 65  
 278 <210> SEQ ID NO: 10  
 279 <211> LENGTH: 8

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/936,702

DATE: 10/04/2001

TIME: 14:41:19

Input Set : A:\60771 seq.txt

Output Set: N:\CRF3\10042001\I936702.raw

280 <212> TYPE: PRT  
281 <213> ORGANISM: Artificial Sequence  
283 <220> FEATURE:  
284 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide,✓  
285 encoded for by SEQ ID NO: 9  
287 <400> SEQUENCE: 10  
288 Ser Gly His His His His His His  
289 1 5  
292 <210> SEQ ID NO: 11  
293 <211> LENGTH: 131  
294 <212> TYPE: DNA  
295 <213> ORGANISM: Artificial Sequence  
297 <220> FEATURE:  
298 <223> OTHER INFORMATION: Description of Artificial Sequence: reverse✓  
299 oligonucleotide  
301 <400> SEQUENCE: 11  
302 ctatgtactc acttaacttag tggtacacctc gccacacctgag cggccacccccc cgcttccacc 60  
303 ccccccggca ccctccggctt ccggatccac ctccgcctga accggctccccc cactaccggcc 120  
304 acctccggag g 131

5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1009  
10010  
10011  
10012  
10013  
10014  
10015  
10016  
10017  
10018  
10019  
10020  
10021  
10022  
10023  
10024  
10025  
10026  
10027  
10028  
10029  
10030  
10031  
10032  
10033  
10034  
10035  
10036  
10037  
10038  
10039  
10039  
10040  
10041  
10042  
10043  
10044  
10045  
10046  
10047  
10048  
10049  
10049  
10050  
10051  
10052  
10053  
10054  
10055  
10056  
10057  
10058  
10059  
10059  
10060  
10061  
10062  
10063  
10064  
10065  
10066  
10067  
10068  
10069  
10069  
10070  
10071  
10072  
10073  
10074  
10075  
10076  
10077  
10078  
10079  
10079  
10080  
10081  
10082  
10083  
10084  
10085  
10086  
10087  
10088  
10089  
10089  
10090  
10091  
10092  
10093  
10094  
10095  
10096  
10097  
10098  
10098  
10099  
10099  
100100  
100101  
100102  
100103  
100104  
100105  
100106  
100107  
100108  
100109  
100110  
100111  
100112  
100113  
100114  
100115  
100116  
100117  
100118  
100119  
100119  
100120  
100121  
100122  
100123  
100124  
100125  
100126  
100127  
100128  
100129  
100129  
100130  
100131  
100132  
100133  
100134  
100135  
100136  
100137  
100138  
100139  
100139  
100140  
100141  
100142  
100143  
100144  
100145  
100146  
100147  
100148  
100149  
100149  
100150  
100151  
100152  
100153  
100154  
100155  
100156  
100157  
100158  
100159  
100159  
100160  
100161  
100162  
100163  
100164  
100165  
100166  
100167  
100168  
100169  
100169  
100170  
100171  
100172  
100173  
100174  
100175  
100176  
100177  
100178  
100179  
100179  
100180  
100181  
100182  
100183  
100184  
100185  
100186  
100187  
100188  
100189  
100189  
100190  
100191  
100192  
100193  
100194  
100195  
100196  
100197  
100198  
100198  
100199  
100199  
100200  
100201  
100202  
100203  
100204  
100205  
100206  
100207  
100208  
100209  
100209  
100210  
100211  
100212  
100213  
100214  
100215  
100216  
100217  
100218  
100219  
100219  
100220  
100221  
100222  
100223  
100224  
100225  
100226  
100227  
100228  
100229  
100229  
100230  
100231  
100232  
100233  
100234  
100235  
100236  
100237  
100238  
100239  
100239  
100240  
100241  
100242  
100243  
100244  
100245  
100246  
100247  
100248  
100249  
100249  
100250  
100251  
100252  
100253  
100254  
100255  
100256  
100257  
100258  
100259  
100259  
100260  
100261  
100262  
100263  
100264  
100265  
100266  
100267  
100268  
100269  
100269  
100270  
100271  
100272  
100273  
100274  
100275  
100276  
100277  
100278  
100279  
100279  
100280  
100281  
100282  
100283  
100284  
100285  
100286  
100287  
100288  
100289  
100289  
100290  
100291  
100292  
100293  
100294  
100295  
100296  
100297  
100298  
100299  
100299  
100300  
100301  
100302  
100303  
100304  
100305  
100306  
100307  
100308  
100309  
100309  
100310  
100311  
100312  
100313  
100314  
100315  
100316  
100317  
100318  
100319  
100319  
100320  
100321  
100322  
100323  
100324  
100325  
100326  
100327  
100328  
100329  
100329  
100330  
100331  
100332  
100333  
100334  
100335  
100336  
100337  
100338  
100339  
100339  
100340  
100341  
100342  
100343  
100344  
100345  
100346  
100347  
100348  
100349  
100349  
100350  
100351  
100352  
100353  
100354  
100355  
100356  
100357  
100358  
100359  
100359  
100360  
100361  
100362  
100363  
100364  
100365  
100366  
100367  
100368  
100369  
100369  
100370  
100371  
100372  
100373  
100374  
100375  
100376  
100377  
100378  
100379  
100379  
100380  
100381  
100382  
100383  
100384  
100385  
100386  
100387  
100388  
100389  
100389  
100390  
100391  
100392  
100393  
100394  
100395  
100396  
100397  
100398  
100399  
100399  
100400  
100401  
100402  
100403  
100404  
100405  
100406  
100407  
100408  
100409  
100409  
100410  
100411  
100412  
100413  
100414  
100415  
100416  
100417  
100418  
100419  
100419  
100420  
100421  
100422  
100423  
100424  
100425  
100426  
100427  
100428  
100429  
100429  
100430  
100431  
100432  
100433  
100434  
100435  
100436  
100437  
100438  
100439  
100439  
100440  
100441  
100442  
100443  
100444  
100445  
100446  
100447  
100448  
100449  
100449  
100450  
100451  
100452  
100453  
100454  
100455  
100456  
100457  
100458  
100459  
100459  
100460  
100461  
100462  
100463  
100464  
100465  
100466  
100467  
100468  
100469  
100469  
100470  
100471  
100472  
100473  
100474  
100475  
100476  
100477  
100478  
100479  
100479  
100480  
100481  
100482  
100483  
100484  
100485  
100486  
100487  
100488  
100489  
100489  
100490  
100491  
100492  
100493  
100494  
100495  
100496  
100497  
100498  
100499  
100499  
100500  
100501  
100502  
100503  
100504  
100505  
100506  
100507  
100508  
100509  
100509  
100510  
100511  
100512  
100513  
100514  
100515  
100516  
100517  
100518  
100519  
100519  
100520  
100521  
100522  
100523  
100524  
100525  
100526  
100527  
100528  
100529  
100529  
100530  
100531  
100532  
100533  
100534  
100535  
100536  
100537  
100538  
100539  
100539  
100540  
100541  
100542  
100543  
100544  
100545  
100546  
100547  
100548  
100549  
100549  
100550  
100551  
100552  
100553  
100554  
100555  
100556  
100557  
100558  
100559  
100559  
100560  
100561  
100562  
100563  
100564  
100565  
100566  
100567  
100568  
100569  
100569  
100570  
100571  
100572  
100573  
100574  
100575  
100576  
100577  
100578  
100579  
100579  
100580  
100581  
100582  
100583  
100584  
100585  
100586  
100587  
100588  
100589  
100589  
100590  
100591  
100592  
100593  
100594  
100595  
100596  
100597  
100598  
100599  
100599  
100600  
100601  
100602  
100603  
100604  
100605  
100606  
100607  
100608  
100609  
100609  
100610  
100611  
100612  
100613  
100614  
100615  
100616  
100617  
100618  
100619  
100619  
100620  
100621  
100622  
100623  
100624  
100625  
100626  
100627  
100628  
100629  
100629  
100630  
100631  
100632  
100633  
100634  
100635  
100636  
100637  
100638  
100639  
100639  
100640  
100641  
100642  
100643  
100644  
100645  
100646  
100647  
100648  
100649  
100649  
100650  
100651  
100652  
100653  
100654  
100655  
100656  
100657  
100658  
100659  
100659  
100660  
100661  
100662  
100663  
100664  
100665  
100666  
100667  
100668  
100669  
100669  
100670  
100671  
100672  
100673  
100674  
100675  
100676  
100677  
100678  
100679  
100679  
100680  
100681  
100682  
100683  
100684  
100685  
100686  
100687  
100688  
100689  
100689  
100690  
100691  
100692  
100693  
100694  
100695  
100696  
100697  
100698  
100699  
100699  
100700  
100701  
100702  
100703  
100704  
100705  
100706  
100707  
100708  
100709  
100709  
100710  
100711  
100712  
100713  
100714  
100715  
100716  
100717  
100718  
100719  
100719  
100720  
100721  
100722  
100723  
100724  
100725  
100726  
100727  
100728  
100729  
100729  
100730  
100731  
100732  
100733  
100734  
100735  
100736  
100737  
100738  
100739  
100739  
100740  
100741  
100742  
100743  
100744  
100745  
100746  
100747  
100748  
100749  
100749  
1

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/936,702

DATE: 10/04/2001

TIME: 14:41:20

Input Set : A:\60771 seq.txt

Output Set: N:\CRF3\10042001\I936702.raw

L:10 M:270 C: Current Application Number differs, Replaced Application Number

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date

TOP SECRET//NOFORN//COMINT//EYESAFE

A NOVEL CHIMERIC PROTEIN FOR PREVENTION AND  
TREATMENT OF HIV INFECTION

**FIELD OF THE INVENTION**

This invention relates to proteins useful in the prevention and treatment of human immunodeficiency virus (HIV) infection. More specifically, it relates to fusion proteins that bind to two sites on a single target protein, especially when one binding domain of the fusion protein binds to an induced site (on the target protein) that is exposed by the binding of the other binding domain of the fusion protein.

10

**BACKGROUND OF THE INVENTION**

Acquired immune deficiency syndrome (AIDS) is a fatal disease of growing prevalence in the modern world. The agent responsible for this disease, human immunodeficiency virus (HIV), was first identified in 1983. HIV is a T-lymphotropic retrovirus that invades and replicates in cells of the immune system, primarily helper T-lymphocytes. The consequent dysfunction in T-lymphocyte-mediated immunity results in an immuno-compromised condition. Patients usually die of associated opportunistic viral, bacterial or fungal infections. A characteristic laboratory finding in AIDS is the decrease in helper T lymphocytes (CD4), and particularly a steady decrease in the ratio of CD4 to suppressor T lymphocytes CD8 as the disease progresses. Virus binding is primarily mediated by interaction of gp120, the external subunit of the HIV envelope glycoprotein (Env) with CD4 protein and various coreceptor molecules (one of several alternative chemokine receptors). These interactions then activate the gp41 transmembrane subunit of the envelope glycoprotein, to cause fusion between the virus and cell membranes. See *Retroviruses*, Coffin *et al.* (eds.) (1997) CSHP, New York, Ch. 11.

25

The humoral immune system is triggered by HIV infection, though it generally does not provide sufficient protection to ward off the infection. Env is the major target of anti-HIV neutralizing antibodies (Wyatt *et al.* *Nature* 393:705-711, 1998). However, Env has evolved so that its relatively invariant neutralizing determinants are protected from the humoral immune system. Antibodies to these regions therefore are generated at a low frequency and their neutralizing activities *in vivo* are generally weak. Certain variable regions (e.g., the V3 loop) are targets for potent neutralizing antibodies, but these are typically restricted to a limited number of HIV-strains (in other words, they are not broadly cross-reactive). For a list of several gp120 antigenic epitopes and consensus definitions of the conserved and variable regions of gp120, see published PCT application PCT/US98/02766 (publication number WO 98/36087) and Coffin *et al.* (eds.) (1997) CSHP, New York, Ch. 12.

30

A neutralizing monoclonal antibody (MAb) with potent and broadly cross-reactive activity would have great potential value in protocols aimed at preventing HIV infection before or immediately after exposure, for example in neonatal transmission, post-exposure prophylaxis, and as

a topical inhibitor. Such a MAb may also be useful in treating chronic infection (D'Souza *et al. J. Infect. Dis.* 175:1056-1062, 1997). However only a handful of MAbs with the desired broadly cross-reactive neutralizing activities have been described. Because of limited potency and cross-reactivity of these molecules, even the three most promising candidates have questionable clinical value (D'Souza *et al.*, 1997).

Extensive efforts are underway to provide immunological or pharmacological approaches to controlling HIV infection (Coffin *et al.*, 1997, Ch. 12). The specific interaction between gp120 and CD4 has been exploited in efforts to provide a possible treatment for HIV infection. See, *e.g.*, U.S. Patent No. 5,817,767; Capon *et al.*, *Nature* 337:525-531, 1989. A soluble fragment of CD4 (sCD4), comprising the first and second domains of this protein (D1D2) has been generated, and this molecule interacts specifically with gp120, essentially serving as a molecular decoy. sCD4 has been shown to block the spread of HIV between cultured cells (Moore *et al.*, *Science* 250:1139-1142, 1990). However, clinical trials with sCD4 were inconclusive as to the effects on human viral load (Schooley *et al.*, *Ann. Internal Med.* 112:247-253, 1990; Kahn *et al.*, *Ann. Internal Med.* 112:254-261, 1990). Subsequent studies indicated that, unlike laboratory-adapted HIV strains, isolates obtained directly from infected patients (primary isolates) are resistant to neutralization by sCD4 (Darr *et al.*, *Proc. Natl. Acad. Sci.* 87:6574-6578, 1990).

In another approach, researchers have generated an antibody-like molecule by fusing the binding portion of CD4 to the constant region (Fc) of a human IgG heavy chain (see, *e.g.*, Capon *et al.*, *Nature* 337:525-531, 1989; and Byrn *et al.*, *Nature* 344:667-670, 1990). This molecule, termed CD4 immunoadhesin, exploits the native functions of immunoglobulin Fc, such as its ability to fix complement, its ability to mediate antibody-dependent cytotoxicity, and its transfer across the placental barrier. There are significant drawbacks to using Fc receptors in association with CD4, because such a construct may be responsible for targeting HIV to Fc-receptor bearing cells (*e.g.* macrophages), and might lead to increased transmission of HIV-1 across the placental barrier.

A complementary recombinant molecule has also been made, wherein the binding portion of CD4 is fused to the Fv region of an antibody directed to human CD3; this "Janusin" molecule may be able to re-target cytotoxic T-lymphocytes onto HIV-infected cells (Traunecker *et al.*, *Embo J.* 10:3655-3659, 1991; Traunecker *et al.*, *Int. J. Cancer: Supp.* 7:51-52, 1992). Janusin has been reported to inhibit HIV-mediated cell fusion when administered *in vitro* with neutralizing antibody to either gp41 or the V3 loop of gp120 (Allaway *et al.*, *AIDS Res. Hum. Retroviruses* 9:581-587, 1993; U.S. Patent No. 5,817,767). This system is inherently complicated and inefficient because multiple molecules must be co-administered to the subject.

This invention is directed to proteins that address key failures of the prior art.

## SUMMARY OF THE INVENTION

The present invention takes advantage of the finding that the neutralizing activities of MAbs against certain highly conserved determinants of the coreceptor-binding region of gp120 are revealed only when CD4 first binds to gp120 (as in an sCD4-activated fusion assay). Although some MAbs to 5 CD4-induced epitopes (e.g., the human MAbs 17b and 48d, Thali *et al.*, *J. Virol.* 67:3978-3988, 1993) are broadly cross-reactive with Envs from diverse HIV genetic subtypes (Clades), these neutralizing epitopes are only briefly exposed *in vivo*, and therefore have provided poor targets for clinically protective antibody binding.

The inventors have overcome these difficulties by creating a fusion protein containing a 10 fragment of CD4 attached via a linker to a human single chain Fv directed against an induced (for example, a CD4-induced) neutralizing epitope on gp120, for instance a coreceptor-binding determinant of gp120. CD4-binding exposes highly conserved gp120 determinants involved in binding to coreceptor; therefore the provided fusion protein will have the properties of a highly potent, broadly cross-reactive neutralizing antibody with high *in vivo* activity and no Fc-mediated undesirable targeting properties. When the fusion protein is substantially derived from human 15 proteins, it has minimal immunogenicity and toxicity in humans. Such an agent has great value in the prevention of infection during or immediately after HIV exposure (mother/infant transmission, post-exposure prophylaxis, topical inhibitor), and also in the treatment of chronic infection.

Accordingly, a first embodiment of the current invention is a neutralizing bispecific fusion 20 protein capable of binding to two sites on a target protein. This protein has two different binding domains, an inducing-binding domain and an induced-binding domain, functionally linked by a peptide linker. Nucleic acid molecules encoding such fusion proteins are further aspects of this invention. Also encompassed in the invention are protein analogs, derivatives, or mimetics of such 25 neutralizing bispecific fusion proteins. The arrangement of the inducing- and induced-binding domains need not be organized in binding sequence; the amino-proximal or carboxy-proximal binding domain of the fusion protein may be either the induced-binding or the inducing-binding domain.

In certain embodiments, the linker of this invention is of such length and secondary 30 structure that the linker allows the second binding domain to be in binding proximity to the induced epitope of the target protein when the first binding domain is bound to the inducing site of the target protein. The linker may for instance be substantially flexible. Linkers of about 25-100 angstroms (Å), or about 15-100 amino acid residues in length, are examples of linkers of a sufficient length to maintain the second binding domain in binding proximity to the induced epitope. Specific examples of linkers will include one or more occurrences of the amino acid sequence represented by SEQ ID 35 NO: 1. For instance, the invention encompasses bispecific fusion proteins wherein the two binding domains are functionally linked by the amino acid sequence represented by SEQ ID NO: 2.

Targets for bispecific fusion proteins according to this invention include viral envelope proteins. For instance, viral envelope proteins from the human immunodeficiency virus (HIV) are targets for the disclosed invention. In a specific embodiment of the invention, the viral envelope protein target is gp120.

5 In further aspects of the invention, the first binding domain is capable of binding to an inducing site on the target protein, thereby exposing an induced epitope. For instance, the first binding domain can be a ligand such as CD4 or fragments thereof. Alternatively, such a first binding domain may be a binding portion of a variable region of an antibody heavy or light chain. The first binding domain may include, for example, an antibody-binding domain, a single-chain Fv (SCFv), or binding fragments thereof.

10 The second binding domain, which is capable of forming a neutralizing complex with an induced epitope of the target protein, may be for example an antibody or fragments thereof, such as the variable region, Fv, Fab or antigen-binding domain of an antibody. Another example of the second binding domain of the fusion protein is an engineered single-chain Fv (SCFv).

15 In some particular examples where HIV gp120 is the target, and the inducing site is the gp120 CD4 binding site, the induced epitope may be a coreceptor-binding determinant of gp120. Accordingly, aspects of this invention include proteins in which the first binding domain binds to gp120 in such a way as to cause a CD4-induced conformational change in the complexed gp120 that exposes the second binding domain. The first binding domain may be derived from a CD4 molecule, and include CD4 and soluble fragments thereof (sCD4, e.g. D1, D1D2 and other such fragments), and proteins that mimic the biological activity of a CD4 molecule in binding to the inducing site of gp120. In another embodiment of the invention, the first domain of the gp120-targeted bispecific fusion protein is derived from a CD4 anti-idiotypic antibody, or antibodies that mimic CD4 in exposing epitopes.

20 25 The second domain of the gp120-targeted bispecific fusion protein, which binds to an epitope induced by binding of the first fusion domain, may be chosen from domains and fragments of proteins that bind to such CD4 induced epitopes. Antibodies directed to the induced epitopes, as well as the HIV coreceptor (e.g. a chemokine receptor), HIV coreceptor mimics, and fragments of HIV coreceptor proteins, are examples of sources for the second binding domain of a gp120-target bispecific fusion protein of this invention. Examples of chemokine receptors with HIV coreceptor activity include CXCR4, CCR5, CCR2B, and CCR3. Neutralizing antibodies, including 17b and 48d, are examples of antibodies. Fusion proteins wherein the second domain is an engineered single chain Fv (SCFv) derived from such a neutralizing antibody are also encompassed.

30 35 A particular embodiment of this invention is a functional recombinant bispecific fusion protein capable of binding to two sites on gp120, wherein the inducing-binding domain is sCD4; the induced-binding domain is SCFv(17b); and these two domains are linked by a linker of a length sufficient to maintain the SCFv(17b) in binding proximity an SCFv(17b) epitope when sCD4 is

bound to gp120. A prototypical bispecific fusion protein has the amino acid sequence shown in SEQ ID NO: 3. Nucleic acid molecules encoding such a fusion protein are also encompassed; the prototypical nucleic acid molecule has the sequence shown in SEQ ID NO: 4. Vectors and cells comprising this nucleic acid molecule are also encompassed in the current invention, as are transgenic plants and animals that express the nucleic acid molecule.

This invention also provides methods for producing functional recombinant bispecific fusion proteins capable of binding two sites on a target protein. Such a protein can be produced in a prokaryotic or eukaryotic cell (e.g., yeast, insect and mammalian cells), for instance by transforming or transfecting such a cell with a recombinant nucleic acid molecule comprising a sequence which encodes a disclosed bispecific fusion protein. Such transformed cells can then be cultured under conditions that cause production of the fusion protein, which is then recovered through protein purification means. The protein can include a molecular tag, such as a six-histidine tag, to facilitate its recovery. In particular embodiments, the protein has a hexa-histidine (hexa-his) tag, and a thrombin cleavage site.

The invention further provides methods for inactivating a target protein, for instance a gp120 protein, by contacting the target protein with a fusion protein according to this invention. Where the target protein is gp120, this method involves contacting gp120 with a gp120-targeted bispecific fusion protein, for instance sCD4-SCFv(17b). Proteins according to the current invention can also be used to neutralize a human immunodeficiency virus, by contacting the human immunodeficiency virus with a gp120-targeted fusion protein according to this invention. Binding of a viral or recombinant gp120 protein to soluble CD4 or lymphocyte CD4 can also be blocked and/or prevented by contacting the gp120 protein with gp120-targeted fusion protein. In any of these methods, a variant protein, analog or mimetic of the fusion protein as provided herein may also be used.

Proteins of the current invention can be used to inhibit virus replication or infectivity in a subject by administering to the subject an amount of the fusion protein (for example the sCD4-SCFv(17b) fusion protein), or a variant protein, analog or mimetic thereof, sufficient to inhibit HIV virus replication or infectivity. The fusion protein can be administered in a pharmaceutical composition, and given therapeutically to a person who is known to be infected with HIV, or prophylactically to help prevent infection in someone who has been exposed to the virus, or is at high risk for exposure. Proteins of this invention can also be administered in combination with another compound for the treatment or prevention of HIV infection, such as an HIV reverse transcriptase (RT), integrase, or protease inhibitor, another HIV-1 neutralizing antibody, or an Env-targeted toxin. The other drug may be an HIV antiviral agent, an HIV anti-infective agent, and/or an immunomodulator, or combinations thereof.

## BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a bar graph illustrating relative HIV-1 Env-mediated fusion, in the presence (+) or absence (-) of soluble CD4, between effector cells expressing Env (Ba-L) and target cells expressing CCR5 (co-receptor), but no CD4 (primary receptor).

Figure 2 is a graph showing that antibody 17b does not inhibit HIV-1 Env-mediated fusion in the conventional assay (open box: CXCR4 and CD4 on target cell), but strongly inhibits cell fusion in the sCD4-activated assay (filled circle: only CXCR4 on target cell, sCD4 provided). Additional experiments indicate that this phenomenon occurs with diverse Envs using either CXCR4 or CCR5, and that 17b has broad cross-reactive activity with Envs from genetically diverse HIV-1 isolates.

Figure 3 is a schematic diagram of the CD4-SCFv(17b) genetic construct. The genetic construct encodes sCD4 (D1D2, plus the native CD4 N-terminal signal sequence), followed by the L1 linker (Gly<sub>4</sub>Ser)<sub>n</sub>, which attaches the 17b SCFv (V<sub>H</sub>) attached to V<sub>L</sub> via the L2 linker (Gly<sub>4</sub>Ser)<sub>n</sub>, followed by the thrombin cleavage site and hexa-his tag. There is a BamH I site in the middle of L1 to facilitate production of constructs of different lengths.

Figure 4 is a drawing of mechanisms of binding of a sCD4-SCFv(17b) to gp120, and the resulting neutralization of HIV Env function (fusion and infectivity).

Figs. 4A, 4B, and 4C depict the proposed interaction of HIV (mediated by gp120) with the cell surface receptor CD4 and co-receptor CCR5, and the beginning of fusion (mediated by gp41). Interaction between gp120 and CD4 (Fig. 4A) causes a change in the conformation of gp120 (Fig. 4B), which enables interaction between gp120 and CCR5 (Fig. 4B). This triggers a conformational change in gp41 (Fig. 4C), and leads to fusion. Antibody (for instance, MAb 17b) binds poorly to the transiently exposed epitope on gp120 (Fig. 4B), and thus results in only weak neutralization of fusion or infection.

Figs. 4D and 4E depict a proposed mechanism of sCD4-SCFV(17b) neutralization of fusion. In the presence of the bispecific chimeric fusion protein, the sCD4 domain can bind to gp120 and induce a conformational change in this protein sufficient to permit binding of the SCFV(17b) (Fig. 4D). This effectively blocks fusion between the HIV and infection and the target cell.

## SEQUENCE LISTING

The nucleic and amino acid sequences listed in the accompanying sequence listing are shown using standard letter abbreviations for nucleotide bases, and three letter code for amino acids. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand.

SEQ ID NO: 1 shows the basic repeat cassette for a linker polypeptide.

SEQ ID NO: 2 shows a seven-repeat polypeptide linker

SEQ ID NO: 3 shows the amino acid sequence of the CD4-SCFv(17b) chimeric protein.

SEQ ID NO: 4 shows the nucleic acid sequence of CD4-SCFv(17b).

SEQ ID NO: 5 shows the pair of synthetic oligonucleotides used to form the second half of the Stu I site near the 3' end of CD4 and to produce an Spe I overhang at the 3' end of an intermediate construct (site to be destroyed upon ligation into pCB-3); the oligonucleotide sequences reconstruct the remainder of the second domain of CD4 (through ser<sub>183</sub>), and encode an amino acid sequence including ala<sub>182</sub>ser<sub>183</sub> of CD4 D2 plus an intermediate 37 residue linker (gly<sub>4</sub>ser)<sub>6</sub>gly<sub>4</sub>thr<sub>2</sub>ser, followed directly by the universal translational termination sequence (UTS).

SEQ ID NO: 6 shows the peptide sequence encoded for by the nucleotide sequences in SEQ ID NO: 5.

SEQ ID NO: 7 shows the forward (5') primer used to amplify and attach the 17b SCFv sequence to the CD4- linker sequence in pCD2. Italicics show the region of the primer that overlaps with 17b.

SEQ ID NO: 8 shows the amino acid sequence encoded by the oligonucleotide primer in SEQ ID NO:7. This sequence includes the GlySer residues at the third (Gly<sub>4</sub>Ser) repeat within L1 (encoded by the BamH1 site, followed by the remaining four (Gly<sub>4</sub>Ser) repeats, followed by the first ten residues of the 17b SCFv (shown in italics).

SEQ ID NO: 9 shows the 3' primer used to amplify and attach the 17b SCFv sequence plus the thrombin cleavage site and the hexa-his tag to the CD4- linker sequence in pCD2.

SEQ ID NO: 10 shows the peptide encoded for by the nucleotide sequence in SEQ ID NO: 9.

## DETAILED DESCRIPTION OF THE INVENTION

### I. Abbreviations and Definitions

#### 25 A. Abbreviations

HIV: human immunodeficiency virus

gp120: the external subunit of the envelope glycoprotein complex of HIV

Env: the envelope glycoprotein complex of HIV

MAb: monoclonal antibody

30 Fv: antibody "fragment variable", the variable region of an antibody

SCFv: single-chain antibody variable region

#### B. Definitions

Unless otherwise noted, technical terms are used according to conventional understanding.

35 Definitions of common terms in molecular biology may be found in Benjamin Lewin, *Genes Y*, Oxford University Press, 1994 (ISBN 0-19-854287-9); Kendrew *et al.* (eds.), *The Encyclopedia of Molecular Biology*, Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); and Robert A. Meyers

(ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

In order to facilitate review of the various embodiments of the invention, the following definitions of terms are provided:

5 **Animal:** Living multi-cellular vertebrate organisms, a category that includes, for example, mammals and birds. The term mammal includes both human and non-human mammals. Similarly, the term "subject" includes both human and veterinary subjects.

10 **Bispecific fusion protein:** Proteins that have at least two domains fused together, each domain comprising a binding region capable of forming a specific complex with a target protein. In general, the two domains are genetically fused together, in that nucleic acid molecules that encode each protein domain are functionally linked together, for instance by a linker oligonucleotide, thereby producing a single fusion-encoding nucleic acid molecule. The translated product of such a fusion-encoding nucleic acid molecule is the bispecific fusion protein.

15 The two binding regions of such a bispecific protein may associate with two different binding determinants or epitopes on a single target molecule. One binding domain may bind first to such a target and thereby induce a conformational change in the target such that the binding of the second binding domain to the target is enabled, facilitated, or otherwise increased in affinity. In such an instance, the domain that binds first to the target can be referred to as the inducing-binding domain, while the domain that binds second is the induced-binding domain. These fusion protein domains need not be organized in binding sequence; the amino-proximal binding domain of the fusion protein may be either the induced-binding or the inducing-binding domain; likewise for the carboxy-proximal binding domain.

20 Bispecific fusion proteins can be further labeled according to the target protein they bind to and neutralize. For instance, a bispecific fusion protein according to the current invention that binds to two specific sites on HIV gp120 protein may be referred to as a gp120-targeted bispecific fusion protein.

25 **CD4:** Cluster of differentiation factor 4, a T-cell surface protein that mediates interaction with the MHC class II molecule. CD4 also serves as the primary receptor site for HIV on T-cells during HIV infection.

30 Molecules that are derived from CD4 include fragments of CD4, generated either by chemical (e.g. enzymatic) digestion or genetic engineering means. Such a fragment may be one or more entire CD4 protein domains (for example, extracellular domains D1, D2, D3, and D4), as defined in the immunological literature, or a portion of one or more of these well-defined domains. For instance, a binding molecule or binding domain derived from CD4 would comprise a sufficient portion of the CD4 protein to mediate specific and functional interaction between the binding fragment and a native or viral binding site of CD4. One such binding fragment includes both the D1

and D2 extracellular domains of CD4 (CD4 D1D2), though smaller fragments may also provide specific and functional CD4-like binding. The gp120-binding site has been mapped to D1 of CD4.

The term "CD4-derived molecules" also encompasses analogs (non-protein organic molecules), derivatives (chemically functionalized protein molecules obtained starting with the disclosed protein sequences) or mimetics (three-dimensionally similar chemicals) of the native CD4 structure, as well as proteins sequence variants or genetic alleles, that maintain the ability to functionally bind to a target molecule.

**CD4-induced conformational change:** A change induced in the three-dimensional conformation of the interacting gp120 protein when CD4 specifically interacts with gp120 to form a complex. One characteristic of such a change is the exposure of at least one induced epitope on the interacting gp120 molecule. An epitope induced by such a change is called a CD4-induced epitope. Such a CD4-induced epitope may for instance include gp120 epitopes at or near the co-receptor-binding region of the protein.

In addition to CD4 binding, the binding of other molecules may induce the exposure of induced epitopes on gp120. Such other inducing molecules are considered CD4-like in terms of their epitope-inducing ability, to the extent that they expose epitopes congruent with or equivalent to those induced epitopes exposed upon the binding of native CD4. These other inducing molecules include, but in no way are limited to, fragments of CD4, for instance sCD4, or a fragment containing the D1 or D1 and D2 domains of native CD4. A mannose-specific lectin (SC) may also serve to expose a CD4-induced epitope (see U.S. Patent No. 5,843,454), as can certain anti-gp120 MAbs.

**Complex (complexed):** Two proteins, or fragments or derivatives thereof, are said to form a complex when they measurably associate with each other in a specific manner. Such association can be measured in any of various ways, both direct and indirect. Direct methods may include co-migration in non-denaturing fractionation conditions, for instance. Indirect measurements of association will depend on secondary effects caused by the association of the two proteins or protein domains. For instance, the formation of a complex between a protein and an antibody may be demonstrated by the antibody-specific inhibition of some function of the target protein. In the case of gp120, the formation of a complex between gp120 and a neutralizing antibody to this protein can be measured by determining the degree to which the antibody inhibits gp120-dependent cell fusion or HIV infectivity. Cell fusion inhibition and infectivity assays are discussed further below.

**Exposing an induced epitope:** The process by which two proteins interact specifically to form a complex (an inducing complex), thereby causing a conformational change in at least one of the two proteins (the target protein) such that at least one previously poorly accessible epitope (an induced epitope) is made accessible to intramolecular interaction. The formation of such an inducing complex will generally cause the exposure of more than one induced epitope, each of which may be thereby rendered accessible for intramolecular interaction.

**HIV coreceptor:** A cell-surface protein other than CD4 involved in the interaction of HIV virus and its subsequent entry into a target cell. These proteins may also be referred to as fusion coreceptors for HIV. Examples of such coreceptor proteins include, for instance, members of the chemokine receptor family (e.g. CXCR4, CCR5, CCR3, and CCR2B).

5 HIV coreceptor proteins interact with coreceptor binding determinants of gp120. In general, it is believed that some of these determinants are exposed on gp120 only after the specific interaction of gp120 with CD4, and the consequent CD4-induced conformational change in the interacting gp120. Thus certain HIV coreceptor binding determinants are, or overlap with, CD4-induced epitopes.

10 Neutralization of gp120 can be achieved by the specific binding of neutralizing proteins or protein fragments or domains to one or more coreceptor binding determinants of gp120, thereby blocking interaction between complexed gp120 and the native coreceptor.

15 **HIV neutralizing ability:** The measurable ability of a molecule to inhibit infectivity of HIV virus, either *in vivo* or *in vitro*. The art is replete with methods for measuring the neutralizing ability of various molecules. Techniques include *in vitro* peripheral blood mononuclear cell (PBMC) based assays (D'Souza *et al.*, 1997); measurement of virion attachment (Mondor *et al.*, *J. Virol.* 72:3623-3634, 1998); neutral red dye uptake and antigen capture assays (U.S. Patent No. 5,695,927); vaccinia-based reporter gene cell fusion assay (Nussbaum *et al.*, *J. Virol.* 68:5411-5422, 1994) (standard and sCD4 activated assays); productive infection assays (measuring gag antigen p24 or RT synthesis) (Karn, *HIV: a practical approach*. Oxford Univ. Press, Cambridge, 1995); and infectivity titer reduction assays (Karn, 1995).

20 In addition, physical interaction between gp120 and CD4 or other CD4-like molecules can be examined by various methods. See, for instance U.S. Patent No. 5,843,454 (measuring conformational changes of gp120 on binding of various proteins by virus release and susceptibility of gp120 to thrombin-mediated cleavage of the V3 loop). Alternately, the ability of the CD4-like molecule to compete for binding to gp120 with either native CD4 or antibody that recognizes the CD4 binding site on gp120 (CD4BS) can be measured. This will allow the calculation of relative binding affinities through standard techniques.

25 The invention also includes analogs, derivatives or mimetics of the disclosed fusion proteins, and which have HIV neutralizing ability. Such molecules can be screened for HIV neutralizing ability by assaying a protein similar to the disclosed fusion protein, in that it has one or more conservative amino acid substitutions, or analogs, derivatives or mimetics thereof, and determining whether the similar protein, analog, derivative or mimetic provides HIV neutralization. The HIV neutralization ability and gp120 binding affinity of these derivative compounds can be measured by any known means, including those discussed in this application

30 **Injectable composition:** A pharmaceutically acceptable fluid composition comprising at least one active ingredient, *e.g.* a bispecific fusion protein. The active ingredient is usually dissolved

or suspended in a physiologically acceptable carrier, and the composition can additionally comprise minor amounts of one or more non-toxic auxiliary substances, such as emulsifying agents, preservatives, and pH buffering agents and the like. Such injectable compositions that are useful for use with the fusion proteins of this invention are conventional; formulations are well known in the art.

**Isolated:** An "isolated" biological component (such as a nucleic acid molecule, protein or organelle) is one that has been substantially separated or purified away from other biological components in the cell of the organism in which the component naturally occurs, *i.e.*, other chromosomal and extra-chromosomal DNA and RNA, proteins and organelles. Nucleic acids and proteins that have been "isolated" include nucleic acids and proteins purified by standard purification methods. The term also embraces nucleic acids and proteins prepared by recombinant expression in a host cell as well as chemically synthesized nucleic acids.

**Neutralizing antibodies:** An antibody that is able to specifically bind to a target protein in such a way as to inhibit the subsequent biological functioning of that target protein is said to be neutralizing of that biological function. In general, any protein that can perform this type of specific blocking activity is considered a neutralizing protein; antibodies are therefore a specific class of neutralizing protein. The complex formed by binding of a neutralizing protein to a target protein is called a neutralizing complex.

Antibodies that bind to viruses and bacteria and thereby prevent the binding of these pathogens to target host cells are said to neutralize the pathogen. Therefore, antibodies that bind to HIV proteins and measurably reduce the ability of the virus to bind to or enter target cells (*e.g.*, T-cells or macrophages) are HIV-neutralizing antibodies. In general, HIV neutralizing antibodies can be broken down into several different classes dependent on what region of the viral envelope protein the antibody binds to. Broad classes of such antibodies include anti-gp41 and anti-gp120 antibodies. There are several antigenic regions on the gp120 protein that provide epitopes for the natural or laboratory generation of HIV neutralizing antibodies (see WO 98/36087). Broadly cross-reactive neutralizing antibodies usually interact with relatively invariant regions of Env.

A primary source of neutralizing antibodies is the peripheral blood of patients infected with the HIV virus. Such primary isolates can be cloned and/or immortalized using standard techniques. In addition to the isolation of naturally-occurring neutralizing antibodies, procedures specifically directed toward their production are known in the art. See U.S. Patent Nos. 5,843,454; 5,695,927; 5,643,756; and 5,013,548 for instance.

**Linker:** A peptide, usually between two and 150 amino acid residues in length that serves to join two protein domains in a multi-domain fusion protein. Examples of specific linkers can be found, for instance, in Hennecke *et al.* (*Protein Eng.* 11:405-410, 1998); and U.S. Patent Nos. 5,767,260 and 5,856,456.

Depending on the domains being joined, and their eventual function in the fusion protein, linkers may be from about two to about 150 amino acids in length, though these limits are given as general guidance only. The tendency of fusion proteins to form specific and non-specific multimeric aggregations is influenced by linker length (Aifthan *et al.*, 1998 *Protein Eng.* 8:725-731, 1998). Thus, shorter linkers will tend to promote multimerization, while longer linkers tend to favor maintenance of monomeric fusion proteins. Aggregation can also be minimized through the use of specific linker sequences, as demonstrated in U.S. Patent No. 5,856,456.

Linkers may be repetitive or non-repetitive. One classical repetitive linker used in the production of single chain Fvs (SCFvs) is the (Gly<sub>4</sub>Ser)<sub>3</sub> (or (GGGGS)<sub>3</sub> or (G<sub>4</sub>S)<sub>3</sub>) linker. More recently, non-repetitive linkers have been produced, and methods for the random generation of such linkers are known (Hennecke *et al.*, *Protein Eng.* 11:405-410, 1998). In addition, linkers may be chosen to have more or less secondary character (e.g. helical character, U.S. Patent No. 5,637,481) depending on the conformation desired in the final fusion protein. The more secondary character a linker possesses, the more constrained the structure of the final fusion protein will be. Therefore, substantially flexible linkers that are substantially lacking in secondary structure allow flexion of the fusion protein at the linker.

A linker is capable of retaining a binding domain of a protein in binding proximity of a target binding site when the linker is of sufficient length and flexibility to allow specific interaction between the binding domain and the target binding site. In the case of the bispecific fusion proteins of this invention, a linker that maintains binding proximity permits the sequential binding with the target of first the inducing-binding domain of the fusion protein, then the induced-binding domain. A linker that maintains the domains of a bispecific fusion protein in binding proximity to a target can be considered an operable or functional linker as relates to such a bispecific fusion protein.

**Oligonucleotide:** A linear polynucleotide sequence of between six and 300 nucleotide bases in length.

**Operably linked:** A first nucleic acid sequence is operably linked with a second nucleic acid sequence when the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences are contiguous and, where necessary to join two protein-coding regions, in the same reading frame.

**ORF (open reading frame):** A series of nucleotide triplets (codons) coding for amino acids without any internal termination codons. These sequences are usually translatable into a peptide.

**Parenteral:** Administered outside of the intestine, e.g., not via the alimentary tract. Generally, parenteral formulations are those that will be administered through any possible mode except ingestion. This term especially refers to injections, whether administered intravenously,

intrathecally, intramuscularly, intraperitoneally, or subcutaneously, and various surface applications including intranasal, intradermal, and topical application, for instance.

5 **Pharmaceutically acceptable carriers:** The pharmaceutically acceptable carriers useful in this invention are conventional. *Remington's Pharmaceutical Sciences*, by E. W. Martin, Mack Publishing Co., Easton, PA, 15th Edition (1975), describes compositions and formulations suitable for pharmaceutical delivery of the fusion proteins herein disclosed.

In general, the nature of the carrier will depend on the particular mode of administration being employed. For instance, parenteral formulations usually comprise injectable fluids that include pharmaceutically and physiologically acceptable fluids such as water, physiological saline, balanced 10 salt solutions, aqueous dextrose, glycerol or the like as a vehicle. For solid compositions (e.g., powder, pill, tablet, or capsule forms), conventional non-toxic solid carriers can include, for example, pharmaceutical grades of mannitol, lactose, starch, or magnesium stearate. In addition to biologically-neutral carriers, pharmaceutical compositions to be administered can contain minor amounts of non-toxic auxiliary substances, such as wetting or emulsifying agents, preservatives, and pH buffering agents and the like, for example sodium acetate or sorbitan monolaurate.

15 **Purified:** The term purified does not require absolute purity; rather, it is intended as a relative term. Thus, for example, a purified fusion protein preparation is one in which the fusion protein is more enriched than the protein is in its generative environment, for instance within a cell or in a biochemical reaction chamber. In some embodiments, a preparation of fusion protein is purified such that the fusion protein represents at least 50% of the total protein content of the preparation.

20 **Recombinant:** A recombinant nucleic acid molecule is one that has a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two otherwise separated segments of sequence. This artificial combination can be accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, 25 e.g., by genetic engineering techniques.

Similarly, a recombinant protein is one encoded for by a recombinant nucleic acid molecule.

25 **Sequence identity:** The similarity between two nucleic acid sequences, or two amino acid sequences is expressed in terms of the similarity between the sequences, otherwise referred to as sequence identity. Sequence identity is frequently measured in terms of percentage identity (or similarity or homology); the higher the percentage, the more similar the two sequences are. Homologs of the bispecific fusion protein will possess a relatively high degree of sequence identity when aligned using standard methods.

30 Methods of alignment of sequences for comparison are well known in the art. Various programs and alignment algorithms are described in: Smith and Waterman (*Adv. Appl. Math.* 2: 482, 1981); Needleman and Wunsch (*J. Mol. Biol.* 48: 443-453, 1970); Pearson and Lipman (*Proc. Natl. Acad. Sci., USA* 85:2444-2448, 1988); Higgins and Sharp (*Gene*, 73:237-244, 1988); Higgins and Sharp (*CABIOS* 5:151-153, 1989); Corpet et al. (*Nuc. Acids Res.* 16: 10881-10890, 1988); Huang et al.

(*Comp. Appl. Biosci.* 8:155-165, 1992); and Pearson *et al.* (*Methods in Molecular Biology* 24: 307-331, 1994). Altschul *et al.* (*Nature Genet.*, 6:119-129, 1994) presents a detailed consideration of sequence alignment methods and homology calculations.

5 The alignment tools ALIGN (Myers and Miller, *CABIOS* 4:11-17, 1989) or LFASTA (Pearson and Lipman, *Proc. Natl. Acad. Sci., USA* 85:2444-2448, 1988) may be used to perform sequence comparisons (Internet Program © 1996, W. R. Pearson and the University of Virginia, "fasta20u63" version 2.0u63, release date December 1996). ALIGN compares entire sequences against one another, while LFASTA compares regions of local similarity. These alignment tools and their respective tutorials are available on the Internet at <http://biology.ncsa.uiuc.edu>.

10 Orthologs of the disclosed bispecific fusion proteins are typically characterized by possession of greater than 75% sequence identity counted over the full-length alignment with the amino acid sequence of bispecific fusion protein using ALIGN set to default parameters.

15 The NCBI Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, *J Mol Biol.* 1990 215:403-410, 1990) is available from several sources, including the National Center for Biotechnology Information (NCBI, Bethesda, MD) and on the Internet, for use in connection with the sequence analysis programs blastp, blastn, blastx, tblastn and tblastx. It can be accessed at <http://www.ncbi.nlm.nih.gov/BLAST/>. A description of how to determine sequence identity using this program is available at [http://www.ncbi.nlm.nih.gov/BLAST/blast\\_help.html](http://www.ncbi.nlm.nih.gov/BLAST/blast_help.html).

20 For comparisons of amino acid sequences of greater than about 30 amino acids, the "Blast 2 sequences" function is employed using the default BLOSUM62 matrix set to default parameters, (gap existence cost of 11, and a per residue gap cost of 1). When aligning short peptides (fewer than around 30 amino acids), the alignment should be performed using the Blast 2 sequences function, employing the PAM30 matrix set to default parameters (open gap 9, extension gap 1 penalties). Proteins with even greater similarity to the reference sequences will show increasing percentage identities when 25 assessed by this method, such as at least 90%, at least 92%, at least 94%, at least 95%, at least 97%, at least 98%, or at least 99% sequence identity. In addition, sequence identity can be compared over the full length of one or both binding domains of the disclosed fusion proteins. In such an instance, percentage identities will be essentially similar to those discussed for full-length sequence identity.

30 When significantly less than the entire sequence is being compared for sequence identity, homologs will typically possess at least 80% sequence identity over short windows of 10-20 amino acids, and may possess sequence identities of at least 85%, at least 90%, at least 95%, or at least 99% depending on their similarity to the reference sequence. Sequence identity over such short windows can be determined using LFASTA; methods are described at <http://biology.ncsa.uiuc.edu>. One of skill in the art will appreciate that these sequence identity ranges are provided for guidance only; it is 35 entirely possible that strongly significant homologs could be obtained that fall outside of the ranges provided. The present invention provides not only the peptide homologs that are described above, but also nucleic acid molecules that encode such homologs.

An alternative indication that two nucleic acid molecules are closely related is that the two molecules hybridize to each other under stringent conditions. Stringent conditions are sequence-dependent and are different under different environmental parameters. Generally, stringent conditions are selected to be about 5° C to 20° C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Conditions for nucleic acid hybridization and calculation of stringencies can be found in Sambrook *et al.* (*In Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989) and Tijssen (*Laboratory Techniques in Biochemistry and Molecular Biology* Part I, Ch. 2, Elsevier, New York, 1993). Nucleic acid molecules that hybridize under stringent conditions to the disclosed bispecific fusion protein sequences will typically hybridize to a probe based on either the entire fusion protein encoding sequence, an entire binding domain, or other selected portions of the encoding sequence under wash conditions of 0.2 x SSC, 0.1% SDS at 65° C.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences, due to the degeneracy of the genetic code. It is understood that changes in nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences, each encoding substantially the same protein.

**Specific binding agent:** An agent that binds substantially only to a defined target. Thus a gp120-specific binding agent binds substantially only the gp120 protein. As used herein, the term "gp120-specific binding agent" includes anti-gp120 antibodies and other agents that bind substantially only to a gp120 protein.

Anti-gp120 antibodies may be produced using standard procedures described in a number of texts, including Harlow and Lane (*Using Antibodies, A Laboratory Manual*, CSHL, New York, 1999, ISBN 0-87969-544-7). In addition, certain techniques may enhance the production of neutralizing antibodies (U.S. Patents No. 5,843,454; 5,695,927; 5,643,756; and 5,013,548). The determination that a particular agent binds substantially only to gp120 protein may readily be made by using or adapting routine procedures. One suitable *in vitro* assay makes use of the Western blotting procedure (described in many standard texts, including Harlow and Lane, 1999). Western blotting may be used to determine that a given protein binding agent, such as an anti-gp120 monoclonal antibody, binds substantially only to the MSG protein. Antibodies to gp120 are well known in the art.

Shorter fragments of antibodies can also serve as specific binding agents. For instance, FAbs, Fvs, and single-chain Fvs (SCFvs) that bind to gp120 would be gp120-specific binding agents.

**Therapeutically effective amount of a bispecific fusion protein:** A quantity of bispecific fusion protein sufficient to achieve a desired effect in a subject being treated. For instance, this can be the amount necessary to inhibit viral proliferation or to measurably neutralize disease organism binding mechanisms. In general, this amount will be sufficient to measurably inhibit virus (e.g. HIV) replication or infectivity.

An effective amount of bispecific fusion protein may be administered in a single dose, or in several doses, for example daily, during a course of treatment. However, the effective amount of fusion protein will be dependent on the fusion protein applied, the subject being treated, the severity and type of the affliction, and the manner of administration of the fusion protein. For example, a therapeutically effective amount of fusion protein can vary from about 0.01 mg/kg body weight to about 1 g/kg body weight.

The fusion proteins disclosed in the present invention have equal application in medical and veterinary settings. Therefore, the general term "subject being treated" is understood to include all animals (e.g. humans, apes, dogs, cats, horses, and cows) that are or may be infected with a virus or other disease-causing microorganism that is susceptible to bispecific fusion protein-mediated neutralization.

**Transformed:** A transformed cell is a cell into which has been introduced a nucleic acid molecule by molecular biology techniques. As used herein, the term transformation encompasses all techniques by which a nucleic acid molecule might be introduced into such a cell, including transfection with viral vectors, transformation with plasmid vectors, and introduction of naked DNA by electroporation, lipofection, and particle gun acceleration.

**Vector:** A nucleic acid molecule as introduced into a host cell, thereby producing a transformed host cell. A vector may include nucleic acid sequences that permit it to replicate in a host cell, such as an origin of replication. A vector may also include one or more selectable marker genes and other genetic elements known in the art.

## II. Construction, Expression, and Purification of Bispecific Fusion Proteins.

### A. Selection of component domains.

This invention provides generally a bispecific fusion protein that binds to two different sites on a target protein. As such, any target protein that has two different binding sites is an example of a target for a bispecific fusion protein. Particular targets include proteins on which one of the two binding sites (the induced-binding site) is exposed/induced by the binding of the fusion protein to a first binding site (the inducing-binding site) on the target. The choice of protein binding domains for incorporation into the disclosed bispecific fusion protein will be dictated by the target protein chosen. The choice of linker will also be influenced by the target protein and binding sites chosen. In general, the linker used in any bispecific fusion will be of a length and secondary character to hold the induced-binding domain within binding proximity of the target protein induced binding site, once the inducing-binding domain of the fusion protein has formed a specific complex with the target.

In certain embodiments, the target protein is an HIV envelope glycoprotein, for instance HIV-1 gp120. In certain of these and other embodiments, the inducing-binding site is the CD4 binding site on gp120. As such, the inducing-binding domain of the disclosed bispecific fusion protein can be a binding fragment of CD4, for instance sCD4. Alternately, any other molecule that

specifically interacts with gp120 in such a way as to expose one or more induced epitopes would also serve as the source of an inducing-binding protein domain. The specific fragments used to construct the fusion protein should be chosen so that the conformation of the final fusion provides functional and inducing binding to gp120; this can be assayed either directly (e.g., affinity measurements) or indirectly (e.g., neutralization assays).

Non-CD4-derived CD4 mimics may also be employed as sources for inducing-binding domains of the disclosed fusion proteins. For instance, a mannose-specific lectin (SC) may serve to induce CD4 induced conformational changes (see U.S. Patent No. 5,843,454). Alternatively, antibodies that bind the CD4-binding site or another epitope of gp120 and thereby induce a CD4-like conformational change on the complexed protein can also be used.

Non-peptide CD4 analogs can also be used in this invention, for instance organic or non-organic structural analog of the gp120-interacting domain(s) of the CD4 molecule.

Induced-binding domains of a gp120-targeted fusion protein will include antibodies (or fragments thereof) that recognize induced epitopes of the complexed gp120. In some embodiments, such antibodies are broadly cross-reactive against diverse HIV-1 isolates. Induced epitopes include all of those referred to as CD4-induced (CD4i) epitopes, and in particular those which overlap coreceptor-binding determinants of gp120. Previously identified neutralizing monoclonal antibodies can be used, and include but are not limited to human monoclonal antibodies 17b, 48d, and CG10.

Likewise, induced binding domains of the disclosed chimeric molecules can be non-peptide molecules, for instance organic or non-organic structural analogs of SCFv(17b).

In addition to antibodies that bind induced epitopes of gp120, other sources for induced-binding domains include fragments of coreceptors that specifically interact with a coreceptor binding domain(s) of gp120.

The construction of a gp120-specific bispecific fusion protein can be aided by review of the X-ray crystallographic structure of the ternary complex containing the gp120 core, a two-domain fragment of CD4 (D1D2), and an FAb from a broadly cross-reactive human MAb (17b) directed against the coreceptor-binding determinants of gp120 (Kwong *et al.*, *Nature* 393:648-659, 1998). Computer-based examination of the structural coordinates of this ternary complex, using FRODO (Jones *et al.*, *Meth. Enzymol.* 115:157-171, 1985; Jones, *J. Appl. Cryst.* 11:268-272, 1978; Pflugrath *et al.* Methods and Applications in Crystallography, pages 407-420, Clarendon Press, Oxford), has revealed choices for constructing the chimeric protein. The shortest distance between free termini of CD4 and the 17b FAb is 56 Å, *i.e.* from the free C-terminus of the D1D2 sCD4 fragment to the N-terminus of the 17b FAb heavy chain. A linker connecting these termini would be essentially free of steric hindrance from CD4 and the N-terminus of the 17b light chain. Possible connections could also be made between the N-terminus of CD4 and the C-termini of the 17b heavy or light chains; such connections would require linkers of about 65 Å and about 86 Å, respectively. In the latter two

connections the linker is required to circumvent other portions of the complex, including the bulky variable loops.

#### B. Assembly.

The construction of chimeric molecules, in particular fusion proteins, from domains of known proteins is well known. In general, a nucleic acid molecule that encodes the desired protein domains are joined using standard genetic engineering techniques to create a single, operably linked fusion oligonucleotide. Molecular biological techniques may be found in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989). Specific examples of genetically engineered multi-domain proteins, especially those based on molecules of the immunoglobulin superfamily, joined by various linkers, can be found in the following patent documents:

- 5 U.S. Patent No. 5,856,456 ("Linker for linked fusion polypeptides");
- 10 U.S. Patent No. 5,696,237 ("Recombinant antibody-toxin fusion protein");
- 15 U.S. Patent No. 5,767,260 ("Antigen-binding fusion proteins");
- 20 U.S. Patent No. 5,587,455 ("Cytotoxic agent against specific virus infection"); and
- WO 98/36087 ("Immunological tolerance to HIV epitopes").

Non-peptide analogs that serve as inducing-binding or induced binding domains of the invention can be linked to the opposite domain of the chimeric molecules using known chemical linking techniques, including chemical cross-linking. Cross-linkers are well known, and examples of molecules used for cross-linking can be found, for instance, in U.S. Patent No. 6,027,890 ("Methods and compositions for enhancing sensitivity in the analysis of biological-based assays").

#### C. Expression.

One skilled in the art will understand that there are myriad ways to express a recombinant protein such that it can subsequently be purified. In general, an expression vector carrying the nucleic acid sequence that encodes the desired protein will be transformed into a microorganism for expression. Such microorganisms can be prokaryotic (bacteria) or eukaryotic (e.g., yeast). One example species of bacteria that can be used is *Escherichia coli* (*E. coli*), which has been used extensively as a laboratory experimental expression system. An eukaryotic expression system can be used where the protein of interest requires eukaryote-specific post-translational modifications such as glycosylation. Also, protein can be expressed using a viral (e.g., vaccinia) based expression system.

Protein can also be expressed in animal cell tissue culture, and such a system can be used where animal-specific protein modifications are desirable or required in the recombinant protein.

The expression vector can include a sequence encoding a targeting peptide, positioned in such a way as to be fused to the coding sequence of the bispecific fusion protein. This allows the bispecific fusion protein to be targeted to specific sub-cellular or extra-cellular locations. Various prokaryotic and eukaryotic targeting peptides, and nucleic acid molecules encoding such, are known.

In a prokaryotic expression system, a signal sequence can be used to secrete the newly synthesized protein. In an eukaryotic expression system, the targeting peptide would specify targeting of the hybrid protein to one or more specific sub-cellular compartments, or to be secreted from the cell, depending on which peptide is chosen. Through the use of an eukaryotic secretion signal sequence, the bispecific fusion protein can be expressed in a transgenic animal (for instance a cow, pig, or sheep) in such a manner that the protein is secreted into the milk of the animal.

Vectors suitable for stable transformation of culturable cells are also well known. Typically, such vectors include a multiple-cloning site suitable for inserting a cloned nucleic acid molecule, such that it will be under the transcriptional control of 5' and 3' regulatory sequences. In addition, transformation vectors include one or more selectable markers; for bacterial transformation this is often an antibiotic resistance gene. Such transformation vectors typically also contain a promoter regulatory region (e.g., a regulatory region controlling inducible or constitutive expression), a transcription initiation start site, a ribosome binding site, an RNA processing signal, and a transcription termination site, each functionally arranged in relation to the multiple-cloning site. For production of large amounts of recombinant proteins, an inducible promoter can be used. This permits selective production of the recombinant protein, and allows both higher levels of production than constitutive promoters, and enables the production of recombinant proteins that may be toxic to the expressing cell if expressed constitutively.

In addition to these general guidelines, protein expression/purification kits are produced commercially. See, for instance, the QIAexpress™ expression system from QIAGEN (Chatsworth, CA) and various expression systems provided by INVITROGEN (Carlsbad, CA). Depending on the details provided by the manufacturers, such kits can be used for production and purification of the disclosed bispecific fusion proteins.

#### D. Purification.

One skilled in the art will understand that there are myriad ways to purify recombinant polypeptides, and such typical methods of protein purification may be used to purify the disclosed bispecific fusion proteins. Such methods include, for instance, protein chromatographic methods including ion exchange, gel filtration, HPLC, monoclonal antibody affinity chromatography and isolation of insoluble protein inclusion bodies after over production. In addition, purification affinity-tags, for instance a six-histidine sequence, may be recombinantly fused to the protein and used to facilitate polypeptide purification. A specific proteolytic site, for instance a thrombin-specific digestion site, can be engineered into the protein between the tag and the fusion itself to facilitate removal of the tag after purification.

Commercially produced protein expression/purification kits provide tailored protocols for the purification of proteins made using each system. See, for instance, the QIAexpress™ expression system from QIAGEN (Chatsworth, CA) and various expression systems provided by INVITROGEN (Carlsbad, CA). Where a commercial kit is employed to produce a bispecific fusion

protein, the manufacturer's purification protocol is a particularly disclosed protocol for purification of that protein. For instance, proteins expressed with an amino-terminal hexa-his tag can be purified by binding to nickel-nitrilotriacetic acid (Ni-NTA) metal affinity chromatography matrix (*The QIAexpressionist*, QIAGEN, 1997).

5        Alternately, the binding specificities of either the first or second binding domains, or both, of the disclosed fusion protein may be exploited to facilitate specific purification of the proteins. One method of performing such specific purification would be column chromatography using column resin to which the target molecule, or an epitope or fragment or domain of the target molecule, has been attached.

10      If the bispecific fusion protein is produced in a secreted form, *e.g.* secreted into the milk of a transgenic animal, purification will be from the secreted fluid. Alternately, purification may be unnecessary if the fusion protein can be applied directly to the subject in the secreted fluid (*e.g.* milk).

### III. Variation of a Bispecific Fusion Protein

#### A. Sequence Variants

15      The binding characteristics and therefore neutralizing activity of the fusion proteins disclosed herein lies not in the precise amino acid sequence, but rather in the three-dimensional structure inherent in the amino acid sequences encoded by the DNA sequences. It is possible to recreate the binding characteristics of any of these proteins or protein domains of this invention by recreating the three-dimensional structure, without necessarily recreating the exact amino acid sequence. This can be achieved by designing a nucleic acid sequence that encodes for the three-dimensional structure, but which differs, for instance by reason of the redundancy of the genetic code. Similarly, the DNA sequence may also be varied, while still producing a functional 20 neutralizing protein.

25      Variant neutralizing bispecific binding proteins include proteins that differ in amino acid sequence from the disclosed sequence, but that share structurally significant sequence homology with any of the provided proteins. Variation can occur in any single domain of the fusion protein (*e.g.* the first or second binding domain, or the linker). Variation can also occur in more than one of such domains in any particular variant protein. Such variants may be produced by manipulating the nucleotide sequence of the, for instance a CD4-SCFv(17b)-encoding sequence, using standard procedures, including site-directed mutagenesis or PCR. The simplest modifications involve the substitution of one or more amino acids for amino acids having similar biochemical properties. These so-called conservative substitutions are likely to have minimal impact on the activity of the 30 resultant protein, especially when made outside of the binding site of each domain. Table 1 shows amino acids that may be substituted for an original amino acid in a protein, and which are regarded as conservative substitutions.

Table 1

	Original Residue	Conservative Substitutions
5	Ala	ser
	Arg	lys
	Asn	gln; his
	Asp	glu
	Cys	ser
	Gln	asn
10	Glu	asp
	Gly	pro
	His	asn; gln
	Ile	leu; val
	Leu	ile; val
15	Lys	arg; gln; glu
	Met	leu; ile
	Phe	met; leu; tyr
	Ser	thr
	Thr	ser
20	Trp	tyr
	Tyr	trp; phe
	Val	ile; leu

More substantial changes in protein structure may be obtained by selecting amino acid substitutions that are less conservative than those listed in Table 1. Such changes include changing residues that differ more significantly in their effect on maintaining polypeptide backbone structure (e.g., sheet or helical conformation) near the substitution, charge or hydrophobicity of the molecule at the target site, or bulk of a specific side chain. The following substitutions are generally expected to produce the greatest changes in protein properties: (a) a hydrophilic residue (e.g., seryl or threonyl) is substituted for (or by) a hydrophobic residue (e.g., leucyl, isoleucyl, phenylalanyl, valyl or alanyl); (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain (e.g., lysyl, arginyl, or histadyl) is substituted for (or by) an electronegative residue (e.g., glutamyl or aspartyl); or (d) a residue having a bulky side chain (e.g., phenylalanine) is substituted for (or by) one lacking a side chain (e.g., glycine).

Variant binding domain or fusion protein-encoding sequences may be produced by standard DNA mutagenesis techniques, for example, M13 primer mutagenesis. Details of these techniques are provided in Sambrook (In *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989), Ch. 15. By the use of such techniques, variants may be created which differ in minor ways from the bispecific fusion protein-encoding sequences disclosed. DNA molecules and 40 nucleotide sequences which are derivatives of those specifically disclosed herein and that differ from those disclosed by the deletion, addition, or substitution of nucleotides while still encoding a protein that binds twice to gp120, thereby neutralizing HIV virus infectivity, are comprehended by this invention. In their most simple form, such variants may differ from the disclosed sequences by

alteration of the coding region to fit the codon usage bias of the particular organism into which the molecule is to be introduced.

Alternatively, the coding region may be altered by taking advantage of the degeneracy of the genetic code to alter the coding sequence such that, while the nucleotide sequence is substantially altered, it nevertheless encodes a protein having an amino acid sequence substantially similar to the disclosed fusion sequences. For example, the 18th amino acid residue of the CD4-SCFv(17b) protein (after cleavage of the N-terminal signal sequence) is alanine. The nucleotide codon triplet GCT encodes this alanine residue. Because of the degeneracy of the genetic code, three other nucleotide codon triplets - (GCG, GCC and GCA) - also code for alanine. Thus, the nucleotide sequence of the disclosed CD4-SCFv(17b) encoding sequence could be changed at this position to any of these three alternative codons without affecting the amino acid composition or characteristics of the encoded protein. Based upon the degeneracy of the genetic code, variant DNA molecules may be derived from the cDNA and gene sequences disclosed herein using standard DNA mutagenesis techniques as described above, or by synthesis of DNA sequences. Thus, this invention also encompasses nucleic acid sequences which encode a neutralizing bispecific fusion protein, but which vary from the disclosed nucleic acid sequences by virtue of the degeneracy of the genetic code.

#### B. Peptide Modifications

The present invention includes biologically active molecules that mimic the action of the bispecific fusion proteins of the present invention, and specifically neutralize HIV Env. The proteins of the invention include synthetic embodiments of naturally-occurring proteins described herein, as well as analogues (non-peptide organic molecules), derivatives (chemically functionalized protein molecules obtained starting with the disclosed peptide sequences) and variants (homologs) of these proteins that specifically bind with and neutralize HIV gp120. Each protein of the invention is comprised of a sequence of amino acids, which may be either L- and/or D- amino acids, naturally occurring and otherwise.

Proteins may be modified by a variety of chemical techniques to produce derivatives having essentially the same activity as the unmodified proteins, and optionally having other desirable properties. For example, carboxylic acid groups of the protein, whether carboxyl-terminal or side chain, may be provided in the form of a salt of a pharmaceutically-acceptable cation or esterified to form a C<sub>1</sub>-C<sub>16</sub> ester, or converted to an amide of formula NR<sub>1</sub>R<sub>2</sub> wherein R<sub>1</sub> and R<sub>2</sub> are each independently H or C<sub>1</sub>-C<sub>16</sub> alkyl, or combined to form a heterocyclic ring, such as a 5- or 6-membered ring. Amino groups of the protein, whether amino-terminal or side chain, may be in the form of a pharmaceutically-acceptable acid addition salt, such as the HCl, HBr, acetic, benzoic, toluene sulfonic, maleic, tartaric and other organic salts, or may be modified to C<sub>1</sub>-C<sub>16</sub> alkyl or dialkyl amino or further converted to an amide.

Hydroxyl groups of the protein side chains may be converted to C<sub>1</sub>-C<sub>16</sub> alkoxy or to a C<sub>1</sub>-C<sub>16</sub> ester using well-recognized techniques. Phenyl and phenolic rings of the protein side chains may be

substituted with one or more halogen atoms, such as fluorine, chlorine, bromine or iodine, or with C<sub>1</sub>-C<sub>16</sub> alkyl, C<sub>1</sub>-C<sub>16</sub> alkoxy, carboxylic acids and esters thereof, or amides of such carboxylic acids.

Methylene groups of the protein side chains can be extended to homologous C<sub>2</sub>-C<sub>4</sub> alkynes. Thiols can be protected with any one of a number of well-recognized protecting groups, such as acetamide groups. Those skilled in the art will also recognize methods for introducing cyclic structures into the proteins of this invention to select and provide conformational constraints to the structure that result in enhanced stability.

It also may be advantageous to introduce one or more disulfide bonds to connect the frameworks of the heavy and light chains in the SCFv domain. This modification often enhances the stability and affinity of SCFvs (Reiter *et al.*, *Protein Engineering* 7:697-704, 1994). Here too, the X-ray crystal structure containing the 17 FAb (Kwong *et al.*, *Nature* 393:648-659, 1998) can be used to assess optimal sites for engineering cysteine residues of the heavy and light chains.

Peptidomimetic and organomimetic embodiments are also within the scope of the present invention, whereby the three-dimensional arrangement of the chemical constituents of such peptido- and organomimetics mimic the three-dimensional arrangement of the protein backbone and component amino acid side chains in the bispecific neutralizing fusion protein, resulting in such peptido- and organomimetics of the proteins of this invention having measurable or enhanced neutralizing ability. For computer modeling applications, a pharmacophore is an idealized, three-dimensional definition of the structural requirements for biological activity. Peptido- and organomimetics can be designed to fit each pharmacophore with current computer modeling software (using computer assisted drug design or CADD). See Walters, "Computer-Assisted Modeling of Drugs", in Klegerman & Groves, eds., 1993, *Pharmaceutical Biotechnology*. Interpharm Press: Buffalo Grove, IL, pp. 165-174 and *Principles of Pharmacology* Munson (ed.) 1995, Ch. 102, for descriptions of techniques used in CADD. Also included within the scope of the invention are mimetics prepared using such techniques that produce neutralizing fusion proteins.

#### C. Domain length variation.

It will be appreciated that the protein domains of the current invention may be combined to produce fusion protein molecules without necessarily splicing the components in the same place. It is believed to be possible to use shorter or longer fragments of each component domain, linked by a functional linker. For instance, any component which is spliced within about 10 amino acid residues of the residue specified, and which still provides a functional binding fragment, comprises about the same domain. However, domains of substantially longer or substantially shorter length can be used. For instance, in certain embodiments, the protein can include a leader sequence plus a four-domain CD4 (D1-D4, amino acid residues 1-372), or just the first domain of CD4 (D1 residues 1-113).

**IV. Activity of Fusion Proteins**

It is important to assess the chemical, physical and biological activity of the disclosed bispecific fusion proteins. Among other uses, such assays permit optimization of the domains chosen, as well as optimization of the length and conformation of the linkers used to connect them. 5 Control molecules should be included in each assay; usually such will include each domain alone, as well as the two domains as separate molecules mixed in the reaction, for instance in a 1:1 molar ratio. In the case of a CD4-SCFv(17b) bispecific fusion protein, such controls would include sCD4 and SCFv(17b), for instance.

**A. Fusion protein affinity for target protein**

10 Fusion protein affinity for the target protein can be determined using various techniques. For instance, co-immunoprecipitation analyses with metabolically labeled proteins can be employed to determine binding of sCD4-SCFv proteins, *e.g.* sCD4-SCFv(17b) to soluble HIV-1 gp120, using anti-gp120 MAbs that do not interfere with CD4 binding (*e.g.* MAb D47 that binds to V3), or polyclonal antibody to the C-terminus of gp120. ELISA can also be used to examine the binding characteristics of each domain of the chimera.

**B. Neutralization assays**

Various assays can be used to measure the ability of the disclosed fusion proteins to inhibit function of the target protein. Individual components of the fusion protein will serve as controls. In general, assays will be specific for the target/fusion protein. For instance, many functional analyses can test the ability of sCD4-SCFv fusions to neutralize the HIV Env. It is particularly advantageous to use Envs from diverse HIV-1 strains to test the breadth of inhibition (neutralizing ability) of each fusion protein for different HIV-1 genetic subtypes and different phenotypes (*i.e.* coreceptor usage). In addition, it is advantageous to test such fusion proteins in the standard and sCD4-activated assays for Env-mediated cell fusion. Known HIV-1 neutralizing MAbs and MAbs against CD4-induced 25 epitopes on gp120 are examples of controls for such experiments. Possible synergistic inhibition with other known broadly cross-reactive neutralizing MAbs should be tested (*e.g.* b12, 2F5, F105, 2G12).

In the case of gp120-targeted fusion proteins, the vaccinia-based reporter gene cell fusion assay may be used to assess fusion inhibition (Nussbaum *et al.*, *J. Virol.* 68:5411-5422, 1994). One 30 population of tissue culture cells (*e.g.* BS-C-1, HeLa, or NIH 3T3) uniformly expressing vaccinia virus-encoded binding and fusion-mediated viral envelope glycoprotein(s) is mixed with another population expressing the corresponding cellular receptor(s). In the case of sCD4-SCFv fusions, where the target protein is HIV-1 gp120, one cell population expresses HIV-1 Env, while the other expresses necessary HIV-1 receptors (*e.g.* CD4 and a chemokine receptor). The cytoplasm of either 35 cell population also contains vaccinia virus-encoded bacteriophage T7 DNA polymerase; the cytoplasm of the other contains a transfected plasmid with the *E. coli lacZ* gene linked to the T7 promoter. Upon mixing of the two populations, cell fusion results in activation of the *lacZ* gene,

through the introduction of the T7 RNA polymerase into proximity with the transfected T7 promoter-lacZ in the cytoplasm of the fused cells. The resultant  $\beta$ -galactosidase ( $\beta$ -gal) activity is proportional to the amount of fusion that occurs, and can be measured by colorimetric assay of detergent cell lysates or *in situ* staining. Cell-fusion neutralizing activity of bispecific fusion proteins is therefore assessed by measuring their inhibition of  $\beta$ -gal production.

The gp120-targeted fusions (e.g. sCD4-SCFv) can also be tested for ability to block HIV-1 infection using single round assays (e.g. using indicator cell lines, Vodicka *et al.*, *Virology* 233:193-198, 1997). Target cells expressing CD4 and a specific coreceptor, and containing the lacZ reporter gene linked to the HIV-1 long terminal repeat (LTR), are infected with specific HIV-1 strains (Vodicka, 1997). Integration of an HIV provirus in these cells leads to production of the viral transactivator, Tat, which then turns on expression of the  $\beta$ -gal gene via interaction with LTR. The activity of sCD4-SCFv is assessed by its inhibition of production of  $\beta$ -gal-positive cells (stained blue with X-gal), which is proportional to its ability to block HIV-1 infection.

#### **V. Incorporation of Bispecific Fusion Proteins into Pharmaceutical Compositions**

Pharmaceutical compositions that comprise at least one bispecific fusion protein as described herein as an active ingredient will normally be formulated with a solid or liquid carrier, depending upon the particular mode of administration chosen. The pharmaceutically acceptable carriers and excipients useful in this invention are conventional. For instance, parenteral formulations usually comprise injectable fluids that are pharmaceutically and physiologically acceptable fluid vehicles such as water, physiological saline, other balanced salt solutions, aqueous dextrose, glycerol or the like. Excipients that can be included are, for instance, other proteins, such as human serum albumin or plasma preparations. If desired, the pharmaceutical composition to be administered may also contain minor amounts of non-toxic auxiliary substances, such as wetting or emulsifying agents, preservatives, and pH buffering agents and the like, for example sodium acetate or sorbitan monolaurate.

Other medicinal and pharmaceutical agents, for instance nucleoside derivatives (e.g. AZT) or protease inhibitors, also may be included. It may also be advantageous to include other fusion inhibitors, for instance one or more neutralizing antibodies.

The dosage form of the pharmaceutical composition will be determined by the mode of administration chosen. For instance, in addition to injectable fluids, topical and oral formulations can be employed. Topical preparations can include eye drops, ointments, sprays and the like. Oral formulations may be liquid (e.g., syrups, solutions or suspensions), or solid (e.g., powders, pills, tablets, or capsules). For solid compositions, conventional non-toxic solid carriers can include pharmaceutical grades of mannitol, lactose, starch, or magnesium stearate. Actual methods of preparing such dosage forms are known, or will be apparent, to those skilled in the art.

The pharmaceutical compositions that comprise bispecific fusion protein may be formulated in unit dosage form, suitable for individual administration of precise dosages. One possible unit dosage contains approximately 100 µg of protein. The amount of active compound administered will be dependent on the subject being treated, the severity of the affliction, and the manner of administration, and is best left to the judgment of the prescribing clinician. Within these bounds, the formulation to be administered will contain a quantity of the active component(s) in an amount effective to achieve the desired effect in the subject being treated.

#### VI. Clinical Use of Bispecific Fusion Proteins

The potent viral-neutralizing activity exhibited by the disclosed bispecific fusion proteins makes them useful for treating viral infections in human and other animal subjects. Possibly susceptible viruses include the immunodeficiency viruses, such as HIV and similar or related viruses in simians and other animals. In addition, other viral or microbial systems that involve the interaction of a first inducing and second induced binding site of a single protein will also be susceptible to neutralization using bispecific fusion proteins of the current invention. The bispecific fusion proteins disclosed herein can also be used in highly sensitive detection or purification of target protein.

The bispecific fusion proteins of this invention may be administered to humans, or other animals on whose cells they are effective, in various manners such as topically, orally, intravenously, intramuscularly, intraperitoneally, intranasally, intradermally, intrathecally, and subcutaneously. The particular mode of administration and the dosage regimen will be selected by the attending clinician, taking into account the particulars of the case (e.g., the subject, the disease, and the disease state involved, and whether the treatment is prophylactic or post-infection). Treatment may involve daily or multi-daily doses of bispecific fusion protein(s) over a period of a few days to months, or even years.

If treatment is through the direct administration of cells expressing the bispecific fusion protein to the subject, such cells (e.g. transgenic pluripotent or hematopoietic stem cells or B cells) may be administered at a dose of between about 10<sup>6</sup> and 10<sup>10</sup> cells, on one or several occasions. The number of cells will depend on the patient, as well as the fusion protein and cells chosen to express the protein.

A general strategy for transferring genes into donor cells is disclosed in U.S. Patent No. 5,529,774, which is incorporated by reference. Generally, a gene encoding a protein having therapeutically desired effects is cloned into a viral expression vector, and that vector is then introduced into the target organism. The virus infects the cells, and produces the protein sequence in vivo, where it has its desired therapeutic effect. See, for example, Zabner *et al.*, *Cell* 75:207-216, 1993. As an alternative to adding the sequences encoding the bispecific fusion protein or a homologous protein to the DNA of a virus, it is also possible to introduce such a gene into the somatic DNA of infected or

uninfected cells, by methods that are well known in the art (Sambrook *et al.*, In *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989). These methods can be used to introduce the herein disclosed fusion proteins to human cells to provide long-term resistance to HIV-1 infection or AIDS. For example, gene therapy can be used to secrete the protein at mucosal surfaces, such as the vaginal, rectal, or oral mucosa.

HIV-1 gp120-targeted bispecific fusion proteins, for instance sCD4-SCFv(17b), are particularly useful in the prevention of infection during or immediately after HIV exposure (e.g., mother/infant transmission, post-exposure prophylaxis, and as a topical inhibitor). In such instances, one or more doses of the bispecific fusion protein are administered before or soon after the triggering event. To prevent or ameliorate mother/infant transmission of viral infection, for instance, it may be beneficial to administer the gp120-targeted bispecific fusion protein to the mother intermittently throughout pregnancy, and/or immediately before or following delivery, and/or directly to the newborn immediately after birth. Post-exposure prophylactic treatments may be particularly beneficial where there has been accidental exposure (for instance, a medically related accidental exposure), including but not limited to a contaminated needle-stick or medical exposure to HIV-1 contaminated blood or other fluid.

The present invention also includes combinations of chimeric bispecific fusion proteins with one or more other agents useful in the treatment of disease, e.g. HIV disease. For example, the compounds of this invention may be administered, whether before or after exposure to the virus, in combination with effective doses of other anti-virals, immunomodulators, anti-infectives, and/or vaccines. The term "administration in combination" refers to both concurrent and sequential administration of the active agents.

Examples of antiviral agents that can be used in combination with the chimeric bispecific fusion proteins of the invention are: AL-721 (from Ethigen of Los Angeles, CA), recombinant human interferon beta (from Triton Biosciences of Alameda, CA), Acemannan (from Carrington Labs of Irving, TX), ganciclovir (from Syntex of Palo alto, CA), didehydrodeoxythymidine or d4T (from Bristol-Myers-Squibb), EL10 (from Elan Corp. of Gainesville, GA), dideoxycytidine or ddC (from Hoffman-LaRoche), Novapren (from Novaferon labs, Inc. of Akron, OH), zidovudine or AZT (from Burroughs Wellcome), ribaririn (from Viratek of Costa Mesa, CA), alpha interferon and acyclovir (from Burroughs Wellcome), Indinavir (from Merck & Co.), 3TC (from Glaxo Wellcome), Ritonavir (from Abbott), Saquinavir (from Hoffmann-LaRoche), and others.

Examples of immuno-modulators that can be used in combination with the chimeric bispecific fusion proteins of the invention are AS-101 (Wyeth-Ayerst Labs.), bropirimine (Upjohn), gamma interferon (Genentech), GM-CSF (Genetics Institute), IL-2 (Cetus or Hoffman-LaRoche), human immune globulin (Cutter Biological), IMREG (from Imreg of New Orleans, La.), SK&F106528, and TNF (Genentech).

Examples of some anti-infectives with which the chimeric bispecific fusion proteins can be used include clindamycin with primaquine (from Upjohn, for the treatment of *Pneumocystis pneumonia*), fluconazole (from Pfizer for the treatment of cryptococcal meningitis or candidiasis), nystatin, pentamidine, trimethoprim-sulfamethoxazole, and many others.

5 The combination therapies are of course not limited to the lists provided in these examples, but includes any composition for the treatment of HIV disease (including treatment of AIDS).

## VII. Kits

10 The chimeric proteins disclosed herein can be supplied in the form of a kit for use in prevention and/or treatment of diseases (e.g., HIV infection and AIDS). In such a kit, a clinically effective amount of one or more of the chimeric bispecific fusion proteins is provided in one or more containers. The chimeric bispecific fusion proteins may be provided suspended in an aqueous solution or as a freeze-dried or lyophilized powder, for instance. In certain embodiments, the chimeric proteins will be provided in the form of a pharmaceutical composition.

15 Kits according to this invention can also include instructions, usually written instructions, to assist the user in treating a disease (e.g., HIV infection or AIDS) with a chimeric bispecific fusion protein. Such instructions can optionally be provided on a computer readable medium.

20 The container(s) in which the protein(s) are supplied can be any conventional container that is capable of holding the supplied form, for instance, microfuge tubes, ampoules, or bottles. In some applications, chimeric proteins may be provided in pre-measured single use amounts in individual, typically disposable, tubes or equivalent containers.

25 The amount of a chimeric bispecific fusion protein supplied in the kit can be any appropriate amount, depending for instance on the market to which the product is directed. For instance, if the kit is adapted for research or clinical use, the amount of each chimeric protein provided would likely be an amount sufficient for several treatments.

Certain kits according to this invention will also include one or more other agents useful in the treatment of disease, e.g. HIV disease. For example, such kits may include one or more effective doses of other anti-virals, immunomodulators, anti-infectives, and/or vaccines.

30

## EXAMPLE 1

### Construction of a CD4-SCFv(17b) encoding sequence

35 A gp120-targeted fusion protein, sCD4-SCFv(17b), is constructed by linking the C-terminus of CD4 (D1D2, 183 amino acid residues) to the N-terminus of the heavy chain of the 17b SCFv, which contains the heavy chain at its N-terminus, linked via its C-terminus to the N-terminus of the light chain (see schematic diagram of the construct, Fig. 3). The 17b SCFv DNA was obtained from R. Wyatt and J. Sodroski, Dana Farber Cancer Institute, Boston, MA. The 17b-MAb producing-hybridoma was obtained from J. Robinson, Tuane University.

Linkers were chosen to have sufficient length and flexibility to connect the desired protein segments without inducing unacceptable torsion. For the SCFv, the 15 amino acid residue sequence (Gly<sub>4</sub>Ser)<sub>3</sub> (designated L2) was chosen, which has been employed successfully for production of SCFvs. This sequence confers excellent flexibility with minimal aggregation. The linker between the C-terminus of CD4 and the N-terminus of the SCFv (designated L1; SEQ ID NO: 2), is seven repeats of the same Gly<sub>4</sub>Ser sequence. Conservative estimates indicate that this 35 amino acid residue linker is sufficiently long to allow CD4 and SCFv to bind simultaneously to their respective binding sites on gp120. A schematic of the genetic construct is shown in Fig. 3. A unique BamH I restriction site has been introduced within L1 to enable the production of constructs with shorter or longer linkers, and especially to provide negative controls (linkers too short, thereby not allowing both the CD4 and SCFv moieties of a single molecule to bind simultaneously to their respective binding sites on gp120).

The starting CD4 plasmid is pCB-3, which contains the full-length CD4 cDNA (including its natural 5' signal sequence) in the vaccinia expression plasmid pSC59 (Broder & Berger, *J. Virol.* 67:913-926, 1993). This plasmid was digested with Stu I, which cuts near the end of the 2nd domain of sCD4, and with Spe I, which cuts within the vector downstream of the CD4 insert and leaves a 5' overhang.

Synthetic oligonucleotides (SEQ ID NO: 11) were annealed together to recapitulate the 5' end of the second half of the Stu I site (CCT) and the next two bases (CC) of the CD4 cDNA, and to produce an Spe I overhang at the 3' end (this site to be destroyed upon ligation into pCB-3). The oligonucleotide sequence reconstructs the remainder of the second domain of CD4 (through ser<sub>183</sub>), and encodes the 37 amino acid intermediate linker (gly<sub>4</sub>ser)<sub>6</sub>gly<sub>4</sub>thr<sub>1</sub>ser, followed directly by the universal translational termination sequence (UTS) (SEQ ID NO: 6). A BamH I site has been deliberately included within the linker near the end of the third (gly<sub>4</sub>ser) repeat, to enable subsequent linkage to the 17b SCFv with the exact L1 sequence, and to enable modification of linker length. The resulting intermediate plasmid is designated pCD1. This construct was confirmed by DNA sequence analysis using standard techniques. To facilitate subsequent procedures, the sCD4-linker sequence was recloned into a pSC59 derivative lacking a BamH I site, forming intermediate plasmid pCD2.

The starting 17b plasmid containing the 17b SCFv cDNA in a plasmid vector (pmt del 0) was donated by Dr. Richard Wyatt (Dana Farber Cancer Institute, Boston, MA). The SCFv cDNA is constructed with the heavy chain at the 5' segment and light chain at the 3' segment, attached via DNA encoding the L2 linker (gly<sub>4</sub>ser)<sub>3</sub>. The 17b SCFv construct has a TPA signal sequence at the 5' end, and sequences corresponding to a thrombin cleavage site and a hexa-his tag (to facilitate purification) at the 3' end, followed by a stop codon. A comparable construct without the thrombin cleavage site and hexa-his tag can also be produced.

PCR technology was used to attach the 17b SCFv sequence to the CD4-linker sequence in pCD2. Suitable primers are represented in SEQ ID NOS: 7 and 9. The forward (5') primer (SEQ ID NO: 7) contains a BamH I site near the 5' end (preceded by an overhang), followed by nucleotides that reconstruct the third (gly,ser) plus four additional (gly,ser) repeats; this is followed by 5 nucleotides exactly corresponding to the start of the 17b heavy chain (excluding the 5' signal sequence, beginning at CAG GTG). The 3' primer (SEQ ID NO: 9) begins with convenient restriction sites for cloning into pCD2 (Spe I and others), followed by nucleotides exactly complementary to the 3' end of the 17b SCFv sequence in pmt del 0 (stop codon, hexa-his tag, and thrombin cleavage site).

10 These primers are used to prime the plasmid vector containing the 17b SCFv sequence in pmt del 0, and the resultant PCR product digested with BamH I plus a restriction enzyme that cleaves at the opposite 3' end (e.g., Spe I). This digested fragment is then force-cloned into pCD2 that has been digested with the same enzymes (BamH I and Spe I). The resulting plasmid (designated herein as pCD3) contains the final sCD4-SCFv(17b) construct (with the thrombin cleavage site and hexa-his tag) downstream from the strong, synthetic early/late vaccinia promoter in pSC59. There are convenient, unique restriction sites on each side of the sCD4-SCFv sequence for possible future cloning steps.

20 The 17b SCFv cDNA (including the 5' signal sequence) also has been excised from the pmt del 0 vector by restriction enzyme digestion or PCR, and cloned into the vaccinia expression plasmid pSC59 to provide a control construct.

#### EXAMPLE 2

##### Expression and Purification of CD4-SCFv(17b) fusion protein

###### A. Expression

25 For small amounts of protein expression, vaccinia expression technology can be used to produce the sCD4-SCFv(17b) (as well as the control 17b SCFv protein). The plasmid containing the construct in the vaccinia expression plasmid pSC59 is used to produce a vaccinia recombinant, using standard technology. For such expression, suitable cells (HeLa, BSC-1, etc.) are infected with the recombinant vaccinia virus; after incubation for 24-36 hours at 37° C, the recombinant protein is 30 present in the culture supernatant. Initial biochemical and functional studies can be done with unfractionated supernatant; where necessary, the sCD4-SCFv protein may be purified (see below). Small scale, initial experiments can be performed with small amounts of material (5-20 micrograms, obtained from 1 - 5 X 10<sup>7</sup> cells). The preparation can be scaled up; for such large-scale production, it may be advantageous to employ higher yield technologies for expression of the recombinant proteins (e.g., baculovirus, yeast, or *E. coli*).

35 Expression of the pCD1 secreted protein product (the first two domains of CD4 through ser<sub>183</sub>, plus the 37 amino acid linker) was analyzed. BSC-1 cells were transfected with pCD1 and

infected with wild type vaccinia virus, then incubated overnight at 37° C. Supernatants were analyzed by Western (immunoblot) analysis, using antibodies against CD4. As expected, the protein encoded by pCD1 migrated slightly more slowly than standard purified two-domain sCD4 (Upjohn- Pharmacia, Kalamazoo, MI).

5 The pCD3 full-length sCD4-SCFv(17b) (sCD4-17b) fusion protein has been expressed and tested similarly, and 17b SCFv domain (as cloned into pSC59) can be examined likewise. The sCD4-17b fusion protein (at least a portion of which is secreted) has the expected molecular size (approximately 55 kD) when analyzed by SDS PAGE and Western blotting. The protein reacted strongly with antibodies against CD4 or the hexa-his tag, confirming the presence of these N-  
10 terminal and C-terminal moieties, as well as the correct reading frame.

#### B. Purification

Expressed fusion protein as constructed above with an amino-terminal hexa-his tag was purified using this molecular tag. The tag enables the specific binding and purification of the fusion protein by binding to nickel-nitrilotriacetic acid (Ni-NTA) metal affinity chromatography matrix (see, for instance, *The QIAXpressionist*, QIAGEN, 1997). A hexa-his tag was used in the present examples.

Alternative purification methods include a combination of HPLC and conventional liquid column chromatography (gel filtration; ion exchanger; isoelectric focusing).

#### C. Primary Characterization

20 In order to test gp120 binding to the 17b domain of the sCD4-17b fusion protein, 96-well ELISA plates were first coated with the 13B8.2 anti-CD4 MAb (Beckman Coulter, Chaska, MN, Catalogue no. IM0398), whose epitope on CD4 overlaps determinants involved in binding to gp120. The plates were then incubated with either the purified sCD4-17b or control buffer. When the 25 chimeric protein was captured this way, the 17b moiety remained available to bind gp120 complexed to sCD4; however the sCD4 moiety could not bind free gp120, since it was captured on the plate by the anti-CD4 MAb that blocks the binding site. The plates were incubated with gp120 (IIIB isolate, Ratner *et al.*, *Nature* 313:277-284, 1985) complexed to sCD4. Binding of gp120 was detected by a polyclonal anti-gp120 antiserum, followed by anti-rabbit IgG conjugated to horseradish peroxidase. The plates were washed and incubated with ABTS substrate, and the oxidized product was 30 quantitated by measuring absorbance at 405 nm. The results indicated specific binding: absorbance values were 0.15 with the sCD4-17b chimeric protein, compared to 0.05 with the control buffer.

For testing functionality of the sCD4 region of the chimeric protein, the ELISA plates were first coated with an anti-His tag MAb (QIAGEN Inc., Valencia, CA, Catalog no. 34670), then 35 incubated with either the purified chimeric protein or control buffer. With the chimeric protein captured in this way, the sCD4 moiety was available to bind free gp120; however the 17b moiety could not bind gp120 that was not complexed to sCD4. The plates were incubated with free gp120, and binding was detected as detailed above. The results indicated specific binding: absorbance

values were 0.46 with sCD4-17b, compared to 0.05 with the control buffer. Thus, the ELISA assays confirmed the expected functional binding properties for each moiety of the chimeric protein: 17b moiety bound the gp120/sCD4 complex, and the CD4 moiety bound free gp120.

5

### EXAMPLE 3

#### HIV-envelope neutralization measurements

##### A. Vaccinia-based reporter gene cell fusion

Env-mediated cell fusion activated by CD4 was measured using the vaccinia-based reporter gene assay (Nussbaum *et al.*, *J. Virol.* 68:5411-5422, 1994). For the experiment shown in Table 2 and Fig. 1, effector cells (HeLa) were transfected with plasmid pGINT7- $\beta$ -gal (lacZ linked to T7 promoter), then infected with vaccinia recombinants encoding either the mutant uncleaved Unc Env or the wildtype (WT) SF162 Env (Broder & Berger, *Proc. Natl. Acad. Sci., USA* 92:9004-9008, 1995). Target cells were created by transfecting NIH 3T3 cells with plasmid pGA9-CKR5, containing the CCR5 cDNA linked to a vaccinia promoter (Alkhatib *et al.*, *Science* 272:1955-1958, 1996), then infecting these cells with wild type vaccinia virus WR. Target cells also carry and express a bacteriophage T7 RNA polymerase. Prior to fusion assays, transfected cells were incubated overnight at 31°C to allow expression of recombinant proteins, then washed.

For each fusion assay, mixtures of effector and target cells ( $1 \times 10^3$  of each cell type per well, duplicate wells) were prepared in the absence or presence of sCD4 (100 nM final). After 2.5 hours at 37°C, cells were lysed with NP-40 and  $\beta$ -gal activity was quantitated using standard procedures (Table 2 and Fig. 1). Relative fusion (specific  $\beta$ -gal activity) was determined from the mean of duplicate samples, and calculated as WT-Unc.

10  
15  
20

Table 2

#### Vaccinia-based reporter gene cell fusion assay using soluble CD4

	Total $\beta$ -gal (Raw data)				Relative fusion (WT - Unc)	
	Unc Env (Control)		WT Env (SF162)			
	duplicates	mean	duplicates	mean		
- sCD4	0.50	0.40	0.45	0.50	0.40	0.45
+ sCD4	0.40	0.50	0.45	6.60	5.20	5.90

This vaccinia-based fusion assay can be used to assess the neutralizing ability of the herein disclosed bispecific fusion proteins. The neutralizing ability of MAb 17b was demonstrated to be dependent on the addition of soluble CD4 as follows (see Table 3 and Fig. 2). Effector cells were created by co-infecting HeLa cells with a vaccinia recombinant encoding HIV-1 Env (LAV) (Broder & Berger, *Proc. Natl. Acad. Sci., USA* 92:9004-9008, 1995), and another encoding T7 RNA polymerase. Target cells were created by co-transfecting NIH 3T3 cells with plasmids pYF1-fusin (Feng *et al.*, *Science* 272:872-877, 1996) encoding CXCR4, and pGINT7- $\beta$ -gal (lacZ linked to the T7

30

promoter). The target cells were then infected with vaccinia viruses vCB-3 (encoding CD4, standard assay) (Broder *et al.*, *Virology* 193:483-491, 1993), or WR (wild type virus, sCD4 assay). As background controls, target cells were transfected with pG1NT7-  $\beta$ -gal only (*i.e.*, no coreceptor). Transfected cells were incubated overnight at 31° C to allow expression of recombinant proteins, then washed. Effector cells were incubated 30 minutes at 37° C with the indicated concentration of MAb 17b (Table 3).

For fusion assays, mixtures were prepared between effector and indicated target cells (2 x 10<sup>5</sup> of each cell type per well, duplicate wells); in the standard assay, target cells expressed CXCR4 and CD4, and no soluble CD4 added; in the sCD4 assay, target cells expressed CXCR4 alone, and soluble CD4 was added (200 nM final). After 2.5 hours at 37° C, cells were lysed and  $\beta$ -gal activity was quantitated. Background control  $\beta$ -gal values (standard assay, 0.6; sCD4 assay, 0.2), obtained with target cells lacking coreceptor, were subtracted to give the data presented in Table 3. Data represent percentage of control (no MAb) for each assay.

**Table 3**  
**MAb-mediated inhibition of fusion assay**

[17b] ( $\mu$ g/ml)	Standard Assay		sCD4 Assay	
	$\beta$ -gal	% control	$\beta$ -gal	% control
none	42.3	100.0	11.89	100.0
0.1	39.5	93.4	13.55	113.9
0.5	43.9	103.8	4.66	39.2
1	39.8	94.1	1.68	14.1
5	50.5	119.4	0	0

The effectiveness of the herein described bispecific fusion proteins for neutralizing fusion is tested in a similar manner, by adding varying amounts of the bispecific fusion protein, *e.g.* sCD4-SCFv(17b), to the above assay. Exogenous sCD4 and SCFv(17b) or other gp120-binding proteins need not be added, though they can be used as controls as above, or to determine relative inhibitory efficiencies compared to the bispecific fusion protein. Using this assay, the effects of media from control cells infected with wild-type vaccinia virus WR, were compared with media from cells infected with the recombinant vaccinia virus encoding sCD4-17b. The relative specific  $\beta$ -galactosidase values were 23.4 with the control media and <1 with sCD4-SCFv media. Thus, the chimeric sCD4-17b protein strongly inhibited HIV-1 Env-mediated cell fusion.

#### EXAMPLE 4

##### Large scale production and analysis of sCD4-17b

To produce large amounts of the sCD4-17b protein, the DNA construct has been transferred to the pET11b plasmid vector (Novagen, Madison, WI, Catalog no. 69437-3), which is suitable for high level inducible expression in *E. coli*. This system involves cloning of target genes under control

of strong bacteriophage T7 transcription signal. Once established in a non-expression host bacterial cell, plasmids are then transferred into expression hosts containing a chromosomal copy of the T7 RNA polymerase gene under lacUV5 control, and expression of the recombinant protein of interest (here, sCD4-17b) is induced by the addition of IPTG. The expressed protein is produced at a very high level, and may constitute more than 50% of the total cell protein in the induced culture within a few hours after induction. Western blot results indicate high level expression of the sCD4-17b from the pET11b plasmid.

The protein produced can be denatured and renatured from inclusion bodies to provide a large quantity of functional sCD4-17b protein. This protein can be used for *in vitro* studies to test inhibition in assays of both Env-mediated cell fusion and HIV infection (p24 production).

In addition, the sCD4-17b protein can be used for *in vivo* studies. One *in vivo* model involves SCID mice reconstituted with human thymus plus liver (Pettocello-Montovani *et al.*, *J.Infect.Dis.* 177:337-346, 1998); this system will be used to test whether sCD4-17b inhibits (and to what extent), or prevents, acute HIV-1 infection. This system has been successfully used to demonstrate potent blocking activities of other anti-HIV agents (*e.g.*, protease inhibitors and reverse transcriptase inhibitors, and Env-targeted toxins) (Pettocello-Montovani *et al.*, *J.Infect.Dis.* 177:337-346, 1998).

A second example of an *in vivo* system for testing sCD4-17b activity involves rhesus macaques challenged with SHIV viruses (recombinant viruses containing SIV gag and pol, plus an HIV envelope; Li *et al.*, *J.Viro.* 69:7061-7071, 1995). This system will be used to test whether the sCD4-17b protein inhibits (and to what extent), or prevents, acute SHIV infection.

The effects of sCD4-17b against chronic infection will also be examined, again using the SCID-hu/HIV-1 mouse system and the macaque/SHIV system.

Both *in vitro* and *in vivo* study systems also will be used to test the potency of sCD4-17b protein when used in combination with other anti-HIV agents (*e.g.*, RT and protease inhibitors or other HIV-1 neutralizing MAbs).

The foregoing examples are provided by way of illustration only. Numerous variations on the biological molecules and methods described above may be employed to make and use bispecific fusion molecules capable of binding to two sites on a single protein, and especially two sites on the HIV envelope protein gp120, and for their use in detection, treatment, and prevention of HIV infection. We claim all such subject matter that falls within the scope and spirit of the following claims.

*SEARCHED*

We claim:

1. A neutralizing bispecific fusion protein capable of binding to two sites on a target protein, comprising a first binding domain capable of binding to an inducing site on the target protein, thereby exposing an induced epitope; a second binding domain capable of forming a neutralizing complex with an induced epitope of the target protein; and a linker connecting the first domain to the second domain, wherein at least one of the first and second binding domains comprises a binding portion of a variable region of an antibody heavy or light chain.
2. A protein according to claim 1, wherein the first binding domain comprises a binding portion of a variable region of an antibody heavy or light chain.
3. A protein according to claim 2, wherein the first binding domain comprises an epitope binding domain of an antibody.
4. A protein according to claim 3, wherein the first domain comprises a single-chain Fv (SCFv).
5. A protein according to claim 3, wherein the antibody binding domain mimics a biological activity of a CD4 molecule in binding to the inducing site and exposing the inducing epitope.
6. A protein according to claim 5, wherein the antibody binding domain is derived from a CD4 anti-idiotypic antibody.
7. A protein according to claim 1, wherein the target protein is a viral envelope protein of a virus.
8. A protein according to claim 7, wherein the virus is human immunodeficiency virus (HIV).
9. A protein according to claim 8, wherein the viral envelope protein is gp120.
10. A protein according to claim 2, wherein the first binding domain is derived from a CD4 molecule.
11. A protein according to claim 10, wherein the first binding domain comprises CD4 D1 or CD4D1D2.
12. A protein according to claim 11, wherein the first binding domain is sCD4.
13. A protein according to claim 1, wherein the second binding domain comprises a binding portion of a variable region of an antibody heavy or light chain.
14. A protein according to claim 13, wherein the second binding domain comprises a binding domain of an antibody.
15. A protein according to claim 14, wherein the second binding domain comprises a single-chain Fv (SCFv).
16. A protein according to claim 15, wherein the SCFv is selected from the group consisting of SCFv17b, SCFv48d and SCFvCG10.
17. A protein according to claim 14, wherein the antibody binding domain is derived from a neutralizing monoclonal antibody.

18. A protein according to claim 17, wherein the neutralizing monoclonal antibody is selected from the group consisting of 17b, 48d, and CG10.
19. A protein according to claim 9, wherein the second binding domain mimics a biological activity of an HIV coreceptor molecule in binding to gp120.
20. A protein according to claim 19, wherein the second binding domain comprises a peptide fragment of an HIV coreceptor.
21. A protein according to claim 20, wherein the HIV coreceptor is a chemokine receptor.
22. A protein according to claim 21, wherein the chemokine receptor is selected from the group consisting of CXCR4, CCR2B, CCR3, and CCR5, CCR8, CCR9, CX<sub>3</sub>CR1, US28, or the chemokine receptor related proteins including STRL33, GPR15, APJ, ChemR23, and BLTR.
23. A protein according to claim 9, wherein the induced epitope comprises at least one coreceptor binding determinant of gp120.
24. A protein according to claim 9, wherein the inducing site is a gp120 CD4 binding site.
25. A protein according to claim 14, wherein the binding domain of the antibody is capable of binding to at least one coreceptor binding determinant of gp120.
26. A protein according to claim 1, wherein the linker maintains the second binding domain in binding proximity to the induced epitope when the first binding domain is bound to the inducing site.
27. A protein according to claim 26, wherein the linker is substantially flexible.
28. A protein according to claim 26, wherein the linker is 15-100 angstroms (Å) long.
29. A protein according to claim 26, wherein the linker is 10-100 amino acid residues in length.
30. A protein according to claim 26, wherein the linker comprises at least one occurrence of an amino acid sequence as represented by SEQ ID NO: 1.
31. A protein according to claim 1, wherein the linker comprises at least one occurrence of an amino acid sequence represented by SEQ ID NO: 1.
32. A protein according to claim 31, wherein the linker comprises an amino acid sequence represented by SEQ ID NO: 2.
33. A functional recombinant bispecific fusion protein capable of binding to two sites on gp120, comprising:
- sCD4;
  - SCFv(17b); and
  - a linker of a length sufficient to maintain the SCFv(17b) in binding proximity an SCFv(17b) epitope when sCD4 is bound to gp120.
34. A protein according to claim 33, wherein the linker has an amino acid sequence as represented by SEQ ID NO: 2
35. An isolated nucleic acid molecule encoding a protein according to claim 34.

36. A nucleic acid molecule according to claim 35, wherein the nucleic acid sequence is represented by SEQ ID NO: 3.
37. A protein encoded for by the nucleic acid molecule according to claim 36.
38. An isolated nucleic acid molecule encoding a protein according to claim 1.
39. The nucleic acid molecule according to claim 38, having nucleic acid sequence SEQ ID NO: 4.
40. A transgenic eukaryotic cell comprising the isolated nucleic acid molecule according to claim 38.
41. A method for producing in a eukaryotic cell a functional recombinant bispecific fusion protein capable of binding two sites on a target protein, comprising the steps of:
- transfected the eukaryotic cell with a recombinant nucleic acid molecule according to claim 38;
  - culturing the transfected eukaryotic cell under conditions that cause production of the protein; and
  - recovering the protein produced by the cultured eukaryotic cell.
42. The method of claim 41, wherein the eukaryotic cell is a mammalian cell.
43. The method of claim 41, wherein recovering the protein comprises:
- identifying the protein by the presence of a molecular tag; and
  - separating the protein having the molecular tag so identified from molecules without the tag, so as to recover the protein produced by the cultured eukaryotic cell.
44. A method for inactivating a gp120 protein, comprising the step of: contacting the gp120 protein with a protein according to claim 9, or a variant protein, analog or mimetic thereof.
45. A method for neutralizing a human immunodeficiency virus, comprising the step of: contacting the human immunodeficiency virus with a protein according to claim 9, or a variant protein, analog or mimetic thereof.
46. A method for blocking and preventing the binding of a viral or recombinant gp120 protein to soluble CD4 or lymphocyte CD4, comprising the step of: contacting the gp120 protein with a protein according to claim 9, or a variant protein, analog or mimetic thereof.
47. A method for inhibiting HIV virus replication or infectivity in a subject, comprising administering to the subject an amount of the protein according to claim 9, or a variant protein, analog or mimetic thereof, sufficient to inhibit HIV virus replication or infectivity.
48. A composition comprising the protein according to claim 1, or a variant protein, analog or mimetic thereof.
49. A pharmaceutical composition comprising the protein according to claim 1, or a variant protein, analog or mimetic thereof, and a pharmaceutically acceptable carrier.

50. The method according to claim 47, wherein the protein is administered in the form of a pharmaceutical composition.
51. A protein analog, derivative, or mimetic of the protein of claim 1.
52. A kit for treatment and/or prevention of HIV infection, comprising a clinically effective dose of the neutralizing bispecific fusion protein of claim 1.
53. The kit of claim 52, further comprising instructions.
54. The kit of claim 53, wherein the instructions include directions for administering at least one dose of the neutralizing bispecific fusion protein to a subject in need of such treatment.
55. The kit of claim 52, wherein the neutralizing bispecific fusion protein is provided in the form of a pharmaceutical composition.

1/4

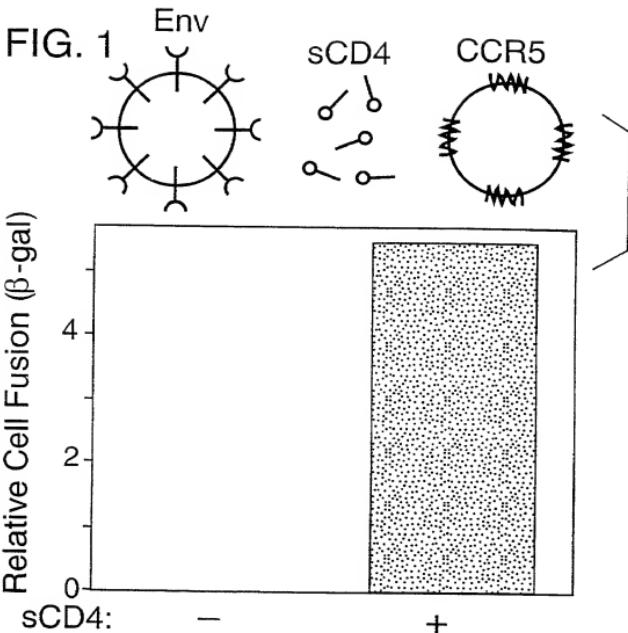
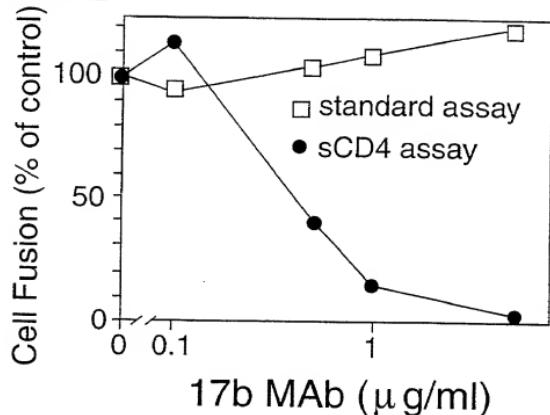


FIG. 2



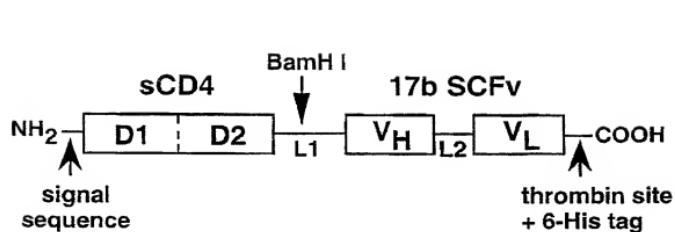
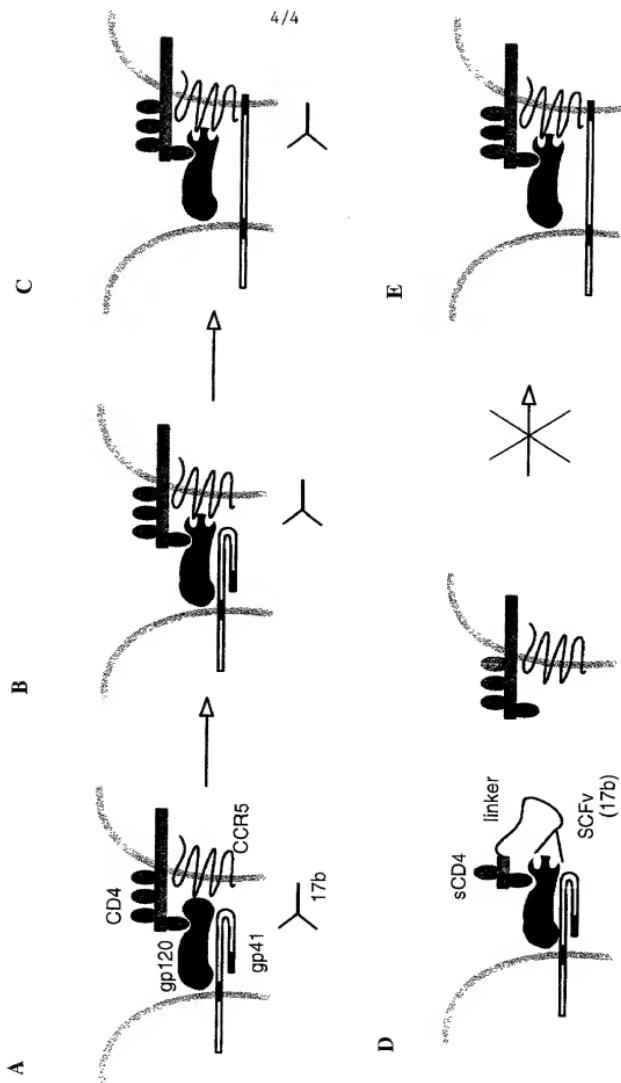
**Figure 3**

Figure 4



PCT/US00/06946 (Application No.)	16 March 2000 (Filing Date)	pending (Status: patented, pending, abandoned)
-------------------------------------	--------------------------------	---

I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application, to file a corresponding international application, and to transact all business in the Patent and Trademark Office connected therewith:

Name	Reg. No.	Name	Reg. No.
James C. Haight	<u>25,588</u>	Jack Spiegel	<u>34,477</u>
Robert Benson	<u>33,612</u>	David R. Sadowski	<u>32,808</u>
Susan S. Rucker	<u>35,762</u>	Steven Ferguson	<u>38,448</u>
John Peter Kim	<u>38,514</u>	Norbert J. Ponzer	<u>40,777</u>
Stephen L. Finley	<u>36,557</u>	Marlene Shinn	<u>46,005</u>
Richard R. Rodriguez	<u>45,980</u>		

with an Associate Power of Attorney to the following:

Name	Reg. No.	Name	Reg. No.
Richard J. Polley	<u>28,107</u>	Lisa M. Caldwell	<u>41,653</u>
William D. Noonan	<u>30,878</u>	Michael D. Jones	<u>41,879</u>
Donald L. Stephens Jr.	<u>34,022</u>	Tanya M. Harding	<u>42,630</u>
Stacey C. Slater	<u>36,011</u>	Gregory L. Maurer	<u>43,781</u>
Susan Alpert Siegel	<u>43,121</u>	Robert F. Scotti	<u>39,830</u>
Wayne W. Rupert	<u>34,420</u>	Sheree L. Kybak	<u>47,913</u>

all of the law firm of Klarquist Sparkman, LLP.

Address all telephone calls to Tanya M. Harding, Ph.D., telephone number 503/226-7391 and facsimile number 503/228-9446.

Address all correspondence to:

**KLARQUIST SPARKMAN, LLP**  
One World Trade Center, Suite 1600  
121 SW Salmon Street  
Portland, OR 97204-2988

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full Name of First Joint Inventor:

Edward A. Berger

Inventor's Signature

Edward A. Berger

8/24/01  
Date

Residence: Rockville, Maryland MD

Citizenship: United States of America

Post Office Address: 5820 Inman Park Circle, Rockville, Maryland 20852

**COMBINED DECLARATION AND POWER OF ATTORNEY  
FOR PATENT APPLICATION**

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am an original, first and joint inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled **A NOVEL CHIMERIC PROTEIN FOR PREVENTION AND TREATMENT OF HIV INFECTION**, the specification of which

- is attached hereto.
- was filed on \_\_\_\_\_ as United States Application No. \_\_\_\_\_.
- was filed on 16 March 2000 as International Application No. PCT/US00/06946.
- and was amended on \_\_\_\_\_ (if applicable).
- with amendments through \_\_\_\_\_ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56. If this is a continuation-in-part application filed under the conditions specified in 35 U.S.C. § 120 which discloses and claims subject matter in addition to that disclosed in the prior copending application, I further acknowledge the duty to disclose material information as defined in 37 CFR § 1.56 which occurred between the filing date of the prior application and the national or PCT international filing date of the continuation-in-part application.

I hereby claim foreign priority benefits under Title 35, United States Code, § 119(a)-(d) of any foreign application(s) for patent or inventor's certificate or of an PCT International application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT International application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) on which priority is claimed:

Prior Foreign Application(s)			Priority Claimed
			<input type="checkbox"/> Yes <input type="checkbox"/> No

I hereby claim the benefit under Title 35, United States Code, § 119(e) of any United States provisional application(s) listed below:

60/124,681	March 16, 1999
(Application No.)	(Filing Date)

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or § 365(c) of any PCT International application(s) designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, § 1.56(a) which occurred between the filing date of the prior application and the national or PCT International filing date of this application:

Full Name of Second Joint Inventor: Christie M. Del Castillo

**Inventor's Signature**

5-039-99/1 08/22/01 2/0 Express Mail Label No. EL754021405US  
Inventor: Christie M. Del Castillo  
Christie M. Del Castillo 8/24/01  
Date

Residence: San Francisco, California

Citizenship: United States of America

Post Office Address: 226 Hugo Street, San Francisco, California 94122  
1222 2nd Ave.

ccc  
8/24/01

I  
SEQUENCE LISTING

&lt;110&gt; Berger, et al.

<120> NOVEL CHIMERIC PROTEIN FOR PREVENTION AND TREATMENT OF  
HIV INFECTION

&lt;130&gt; 4239-54106 Berger PCT

&lt;140&gt;

&lt;141&gt;

&lt;150&gt; 60/124,681

&lt;151&gt; 1999-03-16

&lt;160&gt; 11

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: linker peptide

&lt;400&gt; 1

Gly Gly Gly Gly Ser

1

5

&lt;210&gt; 2

&lt;211&gt; 35

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: seven repeat  
polypeptide linker

&lt;400&gt; 2

Gly Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly

1

5

10

15

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly

20

25

30

Gly Gly Ser

35

&lt;210&gt; 3

&lt;211&gt; 508

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: CD4-scFv(17b)

&lt;400&gt; 3

Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln Leu  
 1 5 10 15

Ala Leu Leu Pro Ala Ala Thr Gln Gly Lys Lys Val Val Leu Gly Lys  
 20 25 30

Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys Ser  
 35 40 45

Ile Gin Phe His Trp Lys Asn Ser Asn Gln Ile Lys Ile Leu Gly Asn  
 50 55 60

Gin Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys Leu Asn Asp Arg Ala  
 65 70 75 80

Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile Ile  
 85 90 95

Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr Ile Cys Glu Val Glu  
 100 105 110

Asp Gln Lys Glu Glu Val Gln Leu Leu Val Phe Gly Leu Thr Ala Asn  
 115 120 125

Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu  
 130 135 140

Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly  
 145 150 155 160

Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu  
 165 170 175

Gln Asp Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys  
 180 185 190

Val Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser  
 195 200 205

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly  
 210 215 220

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly  
 225 230 235 240

Gly Gly Ser Gln Val Gln Leu Leu Gln Ser Gly Ala Glu Val Lys Lys  
 245 250 255

Pro Gly Ser Ser Val Lys Val Ser Cys Lys Ala Ser Gly Asp Thr Phe  
 260 265 270

Ile Arg Tyr Ser Phe Thr Trp Val Arg Gln Ala Pro Gly Gln Gly Leu  
 275 280 285

Glu Trp Met Gly Arg Ile Ile Thr Ile Leu Asp Val Ala His Tyr Ala  
 290 295 300

Pro His Leu Gln Gly Arg Val Thr Ile Thr Ala Asp Lys Ser Thr Ser  
 305 310 315 320

Thr Val Tyr Leu Glu Leu Arg Asn Leu Arg Ser Asp Asp Thr Ala Val

TOKTOK 302560

325	330	335
Tyr Phe Cys Ala Gly Val Tyr Glu Gly Glu Ala Asp Glu Gly Glu Tyr		
340	345	350

Asp Asn Asn Gly Phe Leu Lys His Trp Gly Gln Gly Thr Leu Val Thr		
355	360	365

Val Thr Ser Gly Gly Ser Gly Gly Ser Gly Gly Gly		
370	375	380

Gly Ser Glu Leu Glu Leu Thr Gln Ser Pro Ala Thr Leu Ser Val Ser		
385	390	395

Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Glu Ser Val Ser		
405	410	415

Ser Asp Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu		
420	425	430

Leu Ile Tyr Gly Ala Ser Thr Arg Ala Thr Gly Val Pro Ala Arg Phe		
435	440	445

Ser Gly Ser Gly Ser Gly Ala Glu Phe Thr Leu Thr Ile Ser Ser Leu		
450	455	460

Gln Ser Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Asn Asn Trp		
465	470	475

Pro Pro Arg Tyr Thr Phe Gly Gln Gly Thr Arg Leu Glu Ile Lys Leu		
485	490	495

Val Pro Arg Gly Ser Gly His His His His His His		
500	505	

&lt;210&gt; 4

&lt;211&gt; 1440

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: CD4-scFv(17b)

&lt;400&gt; 4

```

atgaaacccgg gaggcccttt taggcacttg ctctctggtc tgcaactggc gctccctcca 60
gcagccactt aggaaaaaggaa aggtgtgtc ggccaaaaaa gggatacagt gqaactgacc 120
tgcatacgctt cccagaaggaa gagccatacaa ttccacttggaa aaaaactccaa ccagataaaag 180
atctctggatc atcagggtctc ctctttaact aaaggccatc ccaagctggaa tgatcgctg 240
gactcaagaat ggacccctttt ggacccaaaggaa aacttcccccc tgatccatcaaa gaaatcttaag 300
atagaagaact cagatactta catctgtggaa ggccggggacc agaaaggaggg ggtcaattt 360
cttagtggctcg gatgtactgc caactctgtggac accccacccgtc ttccaggggca gacccgtgacc 420
ctgaccctttg agagcccccc ttgttagtggc cocttgcgtc aatgttaggg tccaaagggtt 480
aaaaaacatac agggggggaa gacccttcc ttgttccatgc ttggatctcca ggatgtggc 540
accctggacat gcaactgtttt ggccggggcc aagaagggtgg agttccaaat agacatcggt 600
gtgtctggatcc tcccaaaaggcc ctccggggatc ggccggatgtgg gggggggcc ttccggggca 660
gggtggatccgg gggccggggatc ggccggggcc ggggtggccggg ctccggggcc 720
ggggatccggatcc aggttgcgtatc gtcgtggatcc ggggtggccgg tggatctcc 780
gttgaaaggatc cttgtggatcc accttgcatacata gatatagtttt taatctgggtt 840
cgacaggccc ctggacaagg cttgtggatcc atggggaaaggatc tcatctactt cttgtatgtt 900
gcacactacatc caccggccatc ccggggccatc gtcacgatccatc ccggggacaa gtccacgac 960

```

acagtctacc tggagctgcg gaatctaaga tctgacgata cggccgtata tttctgtgcg 1020  
ggagtgtacg aggaggaggc gacgaggaggaa gaatatgata ataatgggtt tctgaaacat 1080  
tggggccagg gaacccttggc caccgtcacc tcagggtgcg gtggctcccg aggttgtggg 1140  
acggctggcg gggatcttgc attcggatgg acggcgttc cagccacccct gtctgtgtct 1200  
ccaggggaaa gacggccacccct tcctctggcgg gccagtgaga gtgttagtag cgactctagcc 1260  
tggtagccg agaaaaacttgg ccaggctccc aggtctccca tatatgtgc atccacccagg 1320  
gcacccggtg tcccaaggccag gttcagtggc agtgggtctg gggcagaatt cactctacc 1380  
atcaagcagcc tgcagttctga agattttgcg ttgttattact gtcagcgttcaataactgg 1440

<210> 5  
<211> 127  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: synthetic  
oligonucleotide

<400> 5  
cctcccgagg tggcggtagt gggggaggcg gttcaggccg aggtggatcc ggaggccgg 60  
ggtcggccgg ggggtggaaac ggggggtggcg gctctgtgg cggaggatcc actatgttaag 120  
ttagtag 127

<210> 6  
<211> 39  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: peptide  
encoded by SEQ ID NO: 5

<400> 6  
Ala Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly 15  
1 5 10 15  
  
Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser  
20 25 30  
  
Gly Gly Gly Thr Thr Ser  
35

<210> 7  
<211> 103  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: primer

<400> 7  
ttttatggat ccgggtggcg agggtcgggc ggggggtggaa ggggggggtgg cggctccgga 60  
ggcggagggtt cacagggtca gctgctcgag tctggggctg agg 103

<210> 8  
<211> 32  
<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide  
encoded for by SEQ ID NO: 7

<400> 8

Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly  
1 5 10 15  
Ser Gly Gly Gly Ser Gln Val Gln Leu Leu Glu Ser Gly Ala Glu  
20 25 30

<210> 9

<211> 65

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 9

taatttatacg atcacgtgac tagtcctagg cccgggtcaa tggatgtatgat gatgtatggcc 60  
gctgc 65

<210> 10

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide  
encoded for by SEQ ID NO: 9

<400> 10

Ser Gly His His His His His His  
1 5

<210> 11

<211> 131

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: reverse  
oligonucleotide

<400> 11

ctagctactc acttaactag tggtaacctcc gccacactgag ccggccaccccc cgcttccacc 60  
ccccggcccgaa ccctcccgct ccggatccac cttccgcctga accgcctccc cactaccggcc 120  
acctccggag g 131